

```

MPSrch_pp  protein - protein database search, using Smith-Waterman algorithm
Run on:      Thu Jun  8 21:37:37 2000;  MasPar time 14.11 Seconds
Tabular output not generated.  692.177 Million cell updates/sec

```

```
Title: >US-09-316-163-9
Description: (1-207) from US09316163.pep
Perfect Score: 1573 ✓
Sequence: 1 EDCNEIPPRNTEILTGSM.....
```

Scoring table: PAM 150
Gap 11

Searched: 142080 seqs, 47172406 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: pir62

Statistics: Mean 41.851; Variance 66.119; scale 0.633

pred. No. of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

No.	Score	Query	Match	Length	DB	ID	Description	Pred.	No.
1	1573	100.0	449	1	NBHUHS		complement factor H p	0.00e+00	
2	1573	100.0	1231	1	NBHUH		complement factor H p	0.00e+00	
3	1154	73.4	1234	1	NBMHS		complement factor H p	1.13e-21	
4	679	43.2	669	2	S65551		factor H - bovine (fr	5.85e-11	
5	503	32.0	1053	2	S46199		probable complement r	6.86e-99	
6	334	21.2	597	1	S53711		C4Bp alpha chain prec	6.02e-55	
7	310	19.7	560	2	T16833		hyothetical protein	3.60e-44	
8	308	19.6	482	2	JC5092		P-selectin - pig	1.08e-4	
9	308	19.6	830	2	A30359		P-selectin precursor	1.08e-4	
10	303	19.3	283	1	MMV2SP		apolipoprotein H homo	1.69e-44	
11	303	19.3	768	2	A42755		P-selectin precursor	1.69e-44	
12	303	19.3	768	2	I53881		P-selectin - rat	1.69e-44	
13	297	18.9	345	1	NBMS		apolipoprotein H prec	4.55e-44	
14	298	18.9	610	2	A35046		endothelial leukocyte	2.63e-44	
15	298	18.9	610	2	A32606		endothelial leukocyte	2.63e-44	
16	292	18.6	345	1	NBBO		apolipoprotein H prec	7.01e-4	
17	293	18.6	381	1	A26339		decay-accelerating fa	4.06e-4	
18	293	18.6	440	2	A26339		decay-accelerating fa	4.06e-4	
19	290	18.4	551	2	I46709		endothelial leukocyte	2.09e-44	
20	290	18.4	597	1	NBHUC4		C4b-binding protein a	2.09e-44	
21	288	18.3	302	1	MMBE2E		secretory complement	6.22e-44	
22	288	18.3	360	1	MMBE2E		membrane bound comple	6.22e-44	
23	288	18.3	612	2	S53174		endothelial leukocyte	6.22e-44	

287	18.2	551	2	146708	endothelial leukocyte	
25	284	18.1	2	156234	decay-accelerating fa	
26	282	17.9	263	1	C38638	complement control pr
27	281	17.9	349	2	G04913	sperm CDb6 human (f
28	281	17.9	369	2	157938	membrane cofactor pro
29	281	17.9	377	2	154479	membrane cofactor pro
30	281	17.9	384	2	S01896	membrane cofactor pro
31	280	17.8	618	2	B47555	E-selectin precursor
32	279	17.7	656	2	BA5900	complement C3b recept
33	276	17.5	473	2	B38738	coagulation factor C-
34	276	17.5	1019	2	AC38738	coagulation factor C
35	272	17.3	362	2	JC5194	membrane cofactor pro
36	272	17.3	369	2	JC5138	membrane cofactor pro
37	270	17.2	345	1	NBHU	apolipoprotein H prec
38	270	17.2	610	1	146001	C4b-binding protein a
39	270	17.2	1091	1	PL0009	complement C3d/Epstel
40	267	17.0	2039	2	A28507	complement C3b/C4b re
41	267	17.0	2489	2	173012	complement C3b/C4b re
42	266	16.9	485	2	S36772	E-selectin - bovine
43	266	16.9	497	2	JC2054	complement regulatory
44	264	16.8	345	1	JC0465	apolipoprotein H prec
45	265	16.8	469	1	NBMSCA	C4b-binding protein a

ALIGNMENTS

```

RESULT      1
ENTRY
TITLE      NBHUCH      #type complete
ORGANISM    complement factor H precursor, short splice form - human
ORGANISM    complement factor H-related protein: complement protein H
DATE        31-Dec-1993 #sequence_revision 23-Feb-1996 #text_change
           22-Jan-1999
ACCESSIONS  S03013; B60238; A27877; A61103; A26505; S10679
REFERENCE    S00254
#authors     Ripoche, J.; Day, A.-C.; Harris, T.J.B.; Slim, R.B.
#journal      Biochem. J. 1988; 249:593-602
#title        The complete amino acid sequence of human complement factor
              d.
#cross-references MUID:88134059
#accession    S03013
#molecule_type mRNA
#residues     1-449 ##label RIP
#cross-references EMBL:X07523; EMBL:Y00716; NID:g32492;
              PIDN:CA540403.1; PID:g758073
#note         part of this sequence, including the amino end of the
              mature protein was confirmed by protein sequencing
              402-Tyr was also found
REFERENCE     A60238
#authors      Estaller, C.; Schwaeble, W.; Dierich, M.; Weiss, E.H.
#journal      Eur. J. Immunol. (1991) 21:799-802
#title        Human complement factor H: two factor H proteins are derived
              from alternatively spliced transcripts.
#cross-references MUID:91184292
#accession    B60238
#status        not compared with conceptual translation
#molecule_type mRNA
#residues     1-33,434-449 ##label EST
              only portions of this 1.8 kilobase mRNA were sequenced
REFERENCE     A27877
#authors      Schultz, T.F.; Schwaeble, W.; Stanley, K.K.; Weiss, E.;
              Dierich, M.P.
#journal      Eur. J. Immunol. (1986) 16:1351-1355
#title        Human complement factor H: Isolation of cDNA clones and
              partial cDNA sequence of the 38-kDa tryptic fragment
              containing the binding site for C3b.
#cross-references MUID:87054207
#accession    A27877
#molecule_type MRNA
#residues     'IV',55-401,'V',403-449 ##label SCH
              an additional nucleotide present within the codon for
              Glu-310 was found to be a cloning artifact and was
              ignored in translation

```

```

#authors      Schwäbe, W.; Zwirner, J.; Schulz, T.F.; Linke, R.P.;
#journal      Eur. J. Immunol. (1987) 17:1485-1489
#title        Human complement factor H: expression of an additional
#cross-references MIM:88055295
#accession    A61103
#status       not compared with conceptual translation
#molecule_type mRNA
#residues     27-76 ##label SC2
#note         this is a partial sequence of an alternatively spliced
              1.8 kilobase mRNA that is translated to yield a 43 k
              form related to factor H

REFERENCE
#authors      Sim, R.B.; Discipio, R.G.
#journal      Biochem. J. (1982) 205:285-293
#title        Purification and structural studies on the complement-system
              control protein beta-1-H (factor H).
#cross-references MIM:83048213
#accession    A26505
#molecule_type protein
#residues     19-20,'Q',22-29,'V',31-33,'Q',35 ##label SIM
REFERENCE
#authors      Barlow, P.N.; Norman, D.G.; Steinkasserer, A.; Horne, T.J.;
#journal      Biochemistry (1992) 31:3626-3634
#title        Solution structure of the fifth repeat of factor H: A second
              example of the complement control protein module.
#cross-references MIM:92232649
#contents     annotation: NMR structure determination, residues 264-292
#molecule_type protein
#authors      Kristensen, T.; Wetzel, R.A.; Tack, B.F.
#journal      J. Immunol. (1986) 136:3407-3411
#title        Structural analysis of human complement protein H: homology
              with C4b binding protein, beta(2)-glycoprotein I, and the
              Ba fragment of B.
#cross-references MIM:86169701
#accession    S10479
#molecule_type mRNA
#residues     226-401,'Y',403-449 ##label KRI
#cross-references GB:M2383; NID:9180472; PID:AAA52013.1; PID:9180473
COMMENT       Factor H has also been found bound to cell membranes in an unknown
              manner. However, it has at least one cell attachment site motif
              in repeat 4.
COMMENT       Alternative transcripts of 4.3, 1.8, and 1.4 kilobases are expressed
              in liver. See also PIR:NBH0H.
GENETICS
#gene         HFI1
#map_position 19q32-1q32
#cross-references GDB:120041; OMIM:134370
GENETICS
#gene         GDB:HFI1; HF
#map_position 19q32-1q32
#cross-references GDB:129095
#note         the correspondence between the two loci and the sequences
              indicated is unclear; factor H has been reported to have
              several allelic forms

FUNCTION
#description   a cofactor in the inactivation of C3b by serine proteinase I,
              also increases the rate of dissociation of the C3bBb
              complex (C3 convertase) and the (C3b)NB complex (C5
              convertase) in the alternative complement pathway
              #superfamily complement factor H; complement factor H repeat
              homology
              alternative splicing; complement alternate pathway;
              glycoprotein; plasma

KEYWORDS
#pathway
CLASSIFICATION
FEATURE
#domain signal sequence #status predicted #label SIG\
#product complement factor H, short splice form #status
              experimental #label NRT\
#domain complement factor H repeat homology #label FH01V\

```

	85-141	#domain complement factor H repeat homology #label FH02
	146-205	#domain complement factor H repeat homology #label FH03
	210-282	#domain complement factor H repeat homology #label FH04
	246-248	#region cell attachment (R-G-D) motif\
	267-320	#domain complement factor H repeat homology #label FH05
	323-385	#domain complement factor H repeat homology #label FH06
	389-442	#domain complement factor H repeat homology #label FH07
	21-66,52-80,85-129,	
	114-141,146-192,	
	178-205,210-251,	
	237-282,267-309,	
	294-320,325-374,	
	357-385,389-431,	
	416-442	
	217	
SUMMARY	#disulfide_bonds #status predicted\	
	#binding_site carbohydrate (Asn) (covalent) #status	
	absent	
	#length 449 #molecular_weight 51007 #checksum 6077	
Query Match	100.0%; Score 1573; DB 1; Length 449;	
Best Local Similarity 100.0%;	Prod.No. 0.00e+00;	
Matches 207; Conservative 0;	Mismatches 0; Indels 0; Gaps 0;	
Db	19 EDCNCLPPRRNTTEILTGSMSQTYPEGNOALYKCRPEYRSGLGVIWVCKRGWEVALNPRLR 78	
Qy	1 EDCNCLPPRRNTTEILTGSMSDQITPEGIQALYKCRPEYRSGLGVINVCCKRGWEVALNPRLR 60	
Db	79 KCQRPCGHDPEDPFGFTTLTGWNVEFYGVKAVYTCCNEGYOLLGELIYNRCSDTDGTNDI 138	
Qy	61 KCQRPCGHDPEDPFGFTTLTGWNVEFYGVKAVYTCCNEGYOLLGEIYNRCSDTDGTNDI 120	
Db	139 PICEVVKKLPYAPENKGIYSANEPREHFEQAFVNCNSGKIIEGDENHCSDDGFW 198	
Qy	121 PICEVVKKLPYAPENGIYSSANEPREHFQDAVFVNCNSGKIIEGDENHCSDDGFW 180	
Db	199 SKERPKCIVEISCKSPDYINGSPISOKI 225	
Qy	181 SKERPKCIVEISCKSPDYINGSPISOKI 207	
RESULT 2		
ENTRY	NBH0H	#type complete
TITLE	complement factor H precursor, long splice form - human	
ORGANISM	[formal_name Homo sapiens #common_name man	
DATE	31-Dec-1993 #sequence_revision 31-Dec-1993 #ext_change	
	22-Jun-1999	
ACCESSIONS	S00254; A60238; A54726; A61565; A26505; I72654; S66298	
REFERENCE	S00254	
#authors	Ripoche, J.; Day, A.J.; Harris, T.J.R.; Sim, R.B.	
#journal	Biochem. J. (1988) 249:593-602	
#title	The complete amino acid sequence of human complement factor	
	H.	
#cross-references MUID:88134059		
#accession	S00254	
##molecule_type mRNA		
##residues	1-1231 ##label RIP	
##cross-references EMBL:Y00716; NID:931964; PIDN:CMA68704.1; PID:g31965		
#note	402-Tyr was also found	
	parts of this sequence, including the amino and carboxyl	
	ends of the mature protein, were confirmed by protein	
	sequencing	
REFERENCE	A60238	
#authors	Estellier, C.; Schwaebler, W.; Dietrich, M.; Weiss, E.H.	
#journal	Eur. J. Immunol. (1991) 21:799-802	
#title	Human complement factor H: two factor H proteins are derived	
	from alternatively spliced transcripts.	
#cross-references MUID:91184292		
#accession	A60238	
##status	not compared with conceptual translation	
##molecule_type mRNA		
##residues	1-56;1177-1231 ##label EST	
#note	only portions of this 4.3 kilobase mRNA were sequenced	
REFERENCE	A54726	
#authors	Day, A.J.; Ripoche, J.; Lyons, A.; McIntosh, B.; Harris,	

T.J.R.: Sim, R.B.
 Biosci. Rep. (1987) 7:201-207
 #journal Sequence analysis of a cDNA clone encoding the C-terminal end
 #title of human complement factor H.
 #cross-references MUID:88025472
 #accession A54726
 #status not compared with conceptual translation
 #molecule-type mRNA
 #residues 'DEPN', 579-1231 #label DAY
 #cross-references GB:M17517; NID:q180497; PIDN:AAA52016.1; PID:q180498
 #note parts of this sequence were determined by protein sequencing

REFERENCE
 #authors A61565
 #journal Ripoché, J.; Day, A.J.; Willis, A.C.; Belt, K.T.; Campbell, R.D.; Sim, R.B.
 #title Biosci. Rep. (1986) 6:65-72
 #cross-references GB:M17517; NID:q180497; PIDN:AAA52016.1; PID:q180498
 #accession A61565
 #status not compared with conceptual translation
 #molecule-type mRNA
 #residues 'METGRHMLNAKI', 1050-1057, 'T', 1059-1102 #label R12

REFERENCE
 #authors A26505
 #journal Sim, R.B.; Discipio, R.G.
 #title Biochem. J. (1982) 205:285-293
 #cross-references MUID:83048213
 #accession A26505
 #molecule-type protein
 #residues 19-20, 'Q', 22-29, 'V', 31-33, 'Q', 35 #label SIM

REFERENCE
 #authors A4551
 #journal Barlow, P.N.; Norman, D.G.; Steinkasserer, A.; Horne, T.J.; Pearce, J.; Driscoll, P.C.; Sim, R.B.; Campbell, I.D.
 #title Biochemistry (1992) 31:3626-3634
 #cross-references MUID:92232649
 #accession A49224
 #status annotation: NMR structure determination, residues 264-292

REFERENCE
 #authors Norman, D.G.; Barlow, P.N.; Baron, M.; Day, A.J.; Sim, R.B.; Campbell, I.D.
 #journal J. Mol. Biol. (1991) 219:717-725
 #title Three-dimensional structure of a complement control protein module in solution.
 #cross-references MUID:91278097
 #accession I56100
 #status annotation: NMR structure determination, residues 927-985

REFERENCE
 #authors Estélier, C.; Koistinen, V.; Schwaebler, W.; Dierich, M.P.; Weiss, E.H.
 #journal J. Immunol. (1991) 146:3190-3196
 #title Cloning of the 1.4-kb mRNA species of human complement factor H reveals a novel member of the short consensus repeat family related to the carboxy terminal of the classical 150-kDa molecule.
 #cross-references MUID:91201892
 #accession I72654
 #status translated from GB/EMBL/DBJ

REFERENCE
 #molecule-type mRNA
 #residues 1047-1231 #label RES
 #cross-references GB:M65294; NID:q183766; PIDN:AAA35948.1; PID:q183767
 #accession S66298
 #status

REFERENCE
 #authors Caron, J.A.; Bates, R.C.; Smith, A.I.; Tetelz, T.; Arellano, A.; Gordon, D.L.; Burns, G.F.
 #journal Biochim. Biophys. Acta (1996) 1283:305-311
 #title Factor H co-purifies with thrombospondin isolated from platelet secretate.
 #cross-references MUID:96205365
 #accession S66298
 #status preliminary
 #molecule-type protein

##residues 411-419:574-578,580-582 #label CAR
 COMMENT Factor H has also been found bound to cell membranes in an unknown manner. However, it has at least one cell attachment site motif in repeat 4.
 COMMENT Alternative transcripts of 4.3, 1.8, and 1.4 kilobases are expressed in liver. See also PIR:NBH0HS.

GENETICS
 #gene HFI
 #cross-references GDB:120041; OMIM:134370
 #map-position 1q32-1q32
 GENETICS
 #gene HFI
 #cross-references GDB:129095
 #map-position 1q32-1q32
 #note the correspondence between the two loci and the sequences indicated is unclear; factor H has been reported to have several allelic forms

FUNCTION
 #description a cofactor in the inactivation of C3b by serine proteinase I; also increases the rate of dissociation of the C3bd complex (C3 convertase) and the (C3b)nb complex (C5 convertase) in the alternative complement pathway
 #superfamily complement factor H; complement factor H repeat homology
 #alternative splicing: complement alternate pathway: glycoprotein: plasma

KEYWORDS
 #pathway
 CLASSIFICATION
 #product complement factor H, short splice form #status experimental #label MAR
 #domain complement factor H repeat homology #label FH01
 #domain complement factor H repeat homology #label FH02
 #domain complement factor H repeat homology #label FH03
 #domain complement factor H repeat homology #label FH04
 #domain complement factor H repeat homology #label FH05
 #domain complement factor H repeat homology #label FH06
 #domain complement factor H repeat homology #label FH07
 #domain complement factor H repeat homology #label FH08
 #domain complement factor H repeat homology #label FH09
 #domain complement factor H repeat homology #label FH10
 #domain complement factor H repeat homology #label FH11
 #domain complement factor H repeat homology #label FH12
 #domain complement factor H repeat homology #label FH13
 #domain complement factor H repeat homology #label FH14
 #domain complement factor H repeat homology #label FH15
 #domain complement factor H repeat homology #label FH16
 #domain complement factor H repeat homology #label FH17
 #domain complement factor H repeat homology #label FH18
 #domain complement factor H repeat homology #label FH19
 #domain complement factor H repeat homology #label FH20

FEATURE
 1-18
 19-1229
 19-449
 21-80
 85-141
 146-205
 210-262
 246-248
 267-330
 325-385
 389-442
 448-505
 508-554
 569-623
 630-684
 691-744
 753-803
 811-864
 870-926
 931-984
 1048-1102
 1109-1163
 1167-1228
 21-66,52-80,85-129,
 114-141,146-192,
 178-205,210-251,
 237-262,267-309,
 294-320,325-374,
 357-385,389-431,
 416-442,448-494,
 477-505,509-553,
 536-564,569-611,
 597-623,630-673,
 659-684,691-733,
 719-744,753-792,
 781-803,811-853,
 839-864,870-915,
 901-926,931-973,
 959-984,989-1032,
 1018-1043,
 1048-1091,
 1077-1102,

	1109-1152		
	1138-1163		
	1167-1218		
	1201-1228		
	217	#disulfide_bonds #status predicted\	#binding_site carbohydrate (Asn) (covalent) #status absent\
	529,802,822,882,		
	911	#binding_site carbohydrate (Asn) (covalent) #status experimental\	#binding_site carbohydrate (Asn) (covalent) #status predicted
	718,1029,1095	#length 1231	molecular-weight 139124 #checksum 9625
SUMMARY			
Query Match	100.0%	Score 1573:	DB 1: Length 1231:
Best Local Similarity 100.0%;	Pred. No. 0.00e+00;	Mismatches 0;	Indels 0; Gaps 0;
Matches 207;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
Db	19 EDCNELPPRRTEILITGSWSDDQYVEGTQAIVKCRPGYSIGNVIMVCKGEMVALNPLR	78	
Qy	1 EDENELPPRRTEILITGSWSDDQYVEGTQAIVKCRPGYSIGNVIMVCKGEMVALNPLR	60	
Db	79 KCQRPCGHGDPEFGFTLLTGVWVEFGVAAYTCNGCYOLLGGIINRECDTDGWINDI	138	
Qy	61 KCQRPCGHGDPEFGFTLLTGVWVEFGVAAYTCNGCYOLLGGIINRECDTDGWINDI	120	
Db	139 PICEVVKCLPVTAPENGKIYSSAEPDREYHFGQAVRVCNSGXIKEDDEMHCSDDGEFW	198	
Qy	121 PICEVVKCLPVTAPENGKIYSSAEPDREYHFGQAVRVCNSGXIKEDDEMHCSDDGEFW	180	
Db	199 SKRKPCVEISCKSPDYINSPISQKI	225	
Qy	181 SKRKPCVEISCKSPDYINSPISQKI	207	
RESULT	3	NBMSH	#type complete
ENTRY		complement_factor H precursor - mouse	
TITLE		protein beta-1-H	
ALTERNATE_NAMES		#formal_name Mus musculus #common_name house mouse	
ORGANISM		30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 22-Jun-1999	
DATE		A26154; I49711; I49728	
ACCESSIONS		A26154	
REFERENCE		Kristensen, T.; Tack, B.F.	
#authors		Proc. Natl. Acad. Sci. U.S.A. (1986) 83:3963-3967	
#journal		Murine protein H is comprised of 20 repeating units, 61 amino acids in length.	
#title			
#cross-references		MUID:86233353	
#accession		A26154	
##molecule_type		MRNA	
##residues		1-1234 ##label KRI	
REFERENCE		#cross-references GB:M12660; NID:g193724; PIDN:AAA37759.1; PID:g387181	
I49711			
#authors		Natsunume-Sakai, S.; Nonaka, M.; Nonaka, M.; Harada, Y.	
#journal		J. Immunol. (1990) 144:358-362	
#title		Demonstration of an unusual allelic variation of mouse factor H by the complete cDNA sequence of the H.2 allotype.	
#cross-references		MUID:90111033	
#accession		I49711	
##status		preliminary; translated from GB/EMBL/DDBJ	
##molecule_type		MRNA	
##residues		1-18 ##label RES	
REFERENCE		#cross-references GB:M31979; NID:g193726; PIDN:AAA37762.1; PID:g193729	
I49728			
#authors		Munoz-Canooves, P.; Tack, B.F.; Vlk, D.P.	
#journal		Biochemistry (1989) 28:9891-9897	
#title		Analysis of complement factor H mRNA expression: dexamethasone and IFN-gamma increase the level of H in L cells.	
#cross-references		MUID:90148935	
#accession		I49728	
##status		preliminary; translated from GB/EMBL/DBJ	
##molecule_type		MRNA	

```

#####
#residues 1-19 ##label RT2
##cross-references GB:J02891; RID:gl93805; PIDN:AA37795.1; PID:9553926
COMMENT Two codominant alleles of factor H are present in mice.
COMMENT Factor H functions as a cofactor in the inactivation of C3b by
serine proteinase I and also increases the rate of dissociation
of the C3bBb complex (C3 convertase) and the (C3b)nb complex (C5
convertase) in the alternative complement pathway.
GENETICS
#map_position 1
CLASSIFICATION #superfamily complement factor H; complement factor H repeat
KEYWORDS homology
FEATURES complement alternate pathway; duplication; glycoprotein;
plasma
1-18
19-1234
21-80 #domain signal sequence #status predicted #label SIG\
85-141 #product complement factor H #status predicted #label
146-205 MPT\
210-262 #domain complement factor H repeat homology #label FH01\
246-248 #domain complement factor H repeat homology #label FH03\
267-320 #domain complement factor H repeat homology #label FH04\
328-385 #region cell attachment (R-G-D motif)
389-442 #domain complement factor H repeat homology #label FH05\
448-505 #domain complement factor H repeat homology #label FH06\
509-564 #domain complement factor H repeat homology #label FH07\
569-622 #domain complement factor H repeat homology #label FH08\
629-663 #domain complement factor H repeat homology #label FH09\
690-743 #domain complement factor H repeat homology #label FH10\
755-802 #domain complement factor H repeat homology #label FH11\
808-861 #domain complement factor H repeat homology #label FH12\
936-931 #domain complement factor H repeat homology #label FH13\
936-969 #domain complement factor H repeat homology #label FH14\
994-1048 #domain complement factor H repeat homology #label FH15\
1053-1107 #domain complement factor H repeat homology #label FH16\
1114-1168 #domain complement factor H repeat homology #label FH17\
1172-1233 #domain complement factor H repeat homology #label FH18\
21-66, 52-80, 85-129, #domain complement factor H repeat homology #label FH19\
114-141, 146-192,
178-205, 210-251,
237-262, 267-309,
294-320, 325-374,
357-385, 389-431,
416-442, 448-494,
477-505, 509-553,
536-564, 569-610,
597-622, 629-672,
658-683, 690-732,
718-743, 752-791,
780-802, 808-850,
836-861, 867-920,
906-931, 936-978,
964-989, 994-1037,
1023-1048,
1053-1096,
1082-1107,
1114-1157,
1143-1168,
1172-1223,
1206-1233
676,721,773,801,
1030,1061,1225
#disulfide_bonds #status predicted\
#binding_site carbohydrate (asn) (covalent) #status
predicted
#length 1234 #molecular-weight 139081 #checksum 3676
SUMMARY
Query Match 73.4%; Score 1154; DB 1; Length 1234;
Best Local Similarity 68.0%; Pred. No. 1,13e-265;
Matches 140; Conservative 33; Mismatches 33; Indels 0; Gaps 0;
Db 19 EDCKGPPRENSLISWSPOLYPEGQATYKRCQYRTIGTIVYCKNKNWASNSR 78
|||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1 EDCKELPPRNTLTLSWSDQITPEGTQAIYKRCQGRISGNIVYICRKGKEMALNPLR 60
xy

```



```

DATE          31-Jan-1997 #sequence_revision 31-Jan-1997 #text_change
ACCESSIONS    JC5092
REFERENCE      JC5092
#authors      Winkler, H.; Brostjan, C.; Szilmadia, V.; Nataraajan, G.;
               Anrather, J.; Bach, F.H.
#journal       Gene (1996) 176:67-72
#title         The intron-exon structure of the porcine E-selectin-encoding
               gene.
#cross-references MIMD:97075911
#contents      endothelial cells
#accession     JC5092
##molecule-type DNA
##residues     1-482 ##label WIN
#cross-references GB:U37521; NID:g1052974; PID:g1052975
COMMENT        This protein is a member of the selectin family of adhesion
               molecules.
GENETICS
#introns       13/1; 142/1; 178/1; 237/1; 300/1; 363/1; 422/1; 459/1; 466/1
CLASSIFICATION #superfamily complement factor H repeat homology; C-type
                lectin homology
FEATURE
13-139         #domain C-type lectin homology #label LCHN
181-235         #domain complement factor H repeat homology #label FH1\
240-298         #domain complement factor H repeat homology #label FH2\
303-361         #domain complement factor H repeat homology #label FH3\
366-420         #domain complement factor H repeat homology #label FH4
SUMMARY        #length 482 #molecular_weight 52341 #checksum 7729
Query Match    19.6%; Score 308; DB 2; Length 482;
Best Local Similarity 28.9%; Pred. No. 1,08e-47;
Matches 52; Conservative 38; Mismatches 82; Indels 8; Gaps 6;
Db 196 QSLPNTTCAFCCKGFEELIHOCTSSSGMDKRP--TCNAVCDTVGHPRQNDVSC 253
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Qy 22 QYTBETAQVIYCRCGRGISLGNVTMC-RKGWVALNPILKKCRKCRCRGHPGTPTFTFL 80
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 254 NHSTGEFAFYNKSTCFTTCAGFGGLQGPAQT-ECTAOGMTQAQVPCAKVCAPVASOPKNK 312
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Qy 81 TGGNFEEGVXAV-YTCNEGYSQLLEINREDDTG-WTNIDPICEYVKCLPVTAPENG 137
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 313 -LVKRTHSPTGFNYKKSASFCEGEGLGSAQLACTSGOGMTOEVPSCQVVCCSLEY 371
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Qy 138 KIVSSAMPDEHYHFGQVRFCVCSGKIIEGDDEMHCSDDGFWSKEKPCEVISCKSPDV 197
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

RESULT 9
ENTRY          A30359 #type complete
TITLE          P-selectin precursor-human
ALTERNATE_NAMES CD62 antigen; granule membrane protein 140 *
ORGANISM       Homo sapiens #common.name man
DATE           18-Oct-1989 #sequence_revision 30-Sep-1991 #text_change
REFERENCES      29-Aug-1997
A30359
REFERENCE
#authors        Johnston, G.I.; Cook, R.G.; McEver, R.P.
#journal         Cell (1989) 56:1033-1044
#title           Cloning of GMP-140, a granule membrane protein of platelets
               and endothelium: sequence similarity to proteins involved
               in cell adhesion and inflammation.
#cross-references MIMD:89168432
#accession      A30359
##molecule-type mRNA
##residues      1-830 ##label JOH
#cross-references GB:M25322
##note          parts of this sequence, including the amino end of the
               mature protein, were confirmed by protein sequencing
GENETICS
#gene           GDB:SELP; GRMP
#map_position    map_142-1925
CLASSIFICATION #superfamily complement factor H repeat homology; EGF
               homology

```

KEYWORDS

cell adhesion; glycoprotein; phosphatidyl; phosphoprotein; surface antigen; transmembrane protein

FEATURE

1-41 #domain signal sequence #status predicted #label SIG
42-830 #product P-selectin #status experimental #label MAR
163-194 #domain EGF homology #label EGF
200-257 #domain complement factor H repeat homology #label FH01
262-319 #domain complement factor H repeat homology #label FH02
324-381 #domain complement factor H repeat homology #label FH03
386-443 #domain complement factor H repeat homology #label FH04
448-505 #domain complement factor H repeat homology #label FH05
510-567 #domain complement factor H repeat homology #label FH06
572-629 #domain complement factor H repeat homology #label FH07
642-699 #domain complement factor H repeat homology #label FH08
704-761 #domain complement factor H repeat homology #label FH09
772-795 #domain transmembrane #status predicted #label TMN
796-830 #domain intracellular #status predicted #label CIT
54,98,180,212,219,
411,460,518,665,
716,723,741

SUMMARY

#length 830 #molecular-weight 90754 #checksum 2794
Query Match 19.6%; Score 308; DB 2; Length 830;
Best Local Similarity 26.8%; Pred. No. 1.08e-47;
Matches 55; Conservative 44; Mismatches 94; Indels 12; Gaps 12;

Db 199 ECGELELPQHV-LMNCSPHLGFSNOCSEHCITGQVGNPSKIECLASGLWPN-RP-P 255
QY 2 DCNELPBRNRTIELTGSMSDQTYPEGTQAIYKRCRGYRLGAVIWC-RKGWALNLR 60
Db 256 QCLANOC-PELKIPIRGNNICLSAKAPFOHSCSECEGFALVGP-EVVOCTASGYWT 313
QY 61 KQKRCPCGHPGTP-FGRTT-LTGWVFYGYKAYITCEGYQLLGEINRCDPDG-WT 117
Db 314 APAPYKAVQCOHLEAPSEGTM-DC-VHPLTAPVAGSSCKFCQPGYRGVGLDMLRCDIS 371
QY 118 NDPIEYVKKLPVTPENGKIYVSAMEPDREYHFGQAVFVNCNGYKIEGDEMHCSDD 177
Db 372 GHWASAPLPCEALISCEPLSPVHGS 396
QY 178 GFWSKPKPCVCEISCKSPDV-INGS 201

RESULT 10
ENTRY WMYZSP #type complete
TITLE apolipoprotein H homolog precursor - vaccinia virus
ALTERNATE_NAMES 35K secretory protein; C3L protein; vitokine
ORGANISM #formal_name vaccinia virus
#note host Homo sapiens (man)
DATE 31-Dec-1989 #sequence_revision 30-Jun-1990 #text_change 22-Jun-1999

ACCESSIONS A31005; B42504
REFERENCE Kotwal, G.J.; Moss, B.
#authors Nature (1988) 335:176-178
#journal Vaccinia virus encodes a secretory polypeptide structurally related to complement control proteins.

#title
#cross-references MIMD:88318974
#accession A31005
#molecule-type DNA
#residues 1-263 #label KOI
#cross-references GB:X1316; NID:960690; PIDN:CAA31564.1; PID:960691
#experimental_source strain WR
REFERENCE A42501
#authors Goebel, S.J.; Johnson, G.P.; Perkins, M.E.; Davis, S.W.; Winslow, J.P.; Paolletti, E.
#journal Virology (1990) 179:517-563
#title Appendix to "The complete DNA sequence of vaccinia virus".

#accession B42504
#molecule-type DNA
#residues 1-263 #label GOE
#cross-references GB:M35027; NID:9335317; PIDN:AAA7997.1; PID:9335345

#experimental_source strain Copenhagen

REFERENCE

A42531
#authors Goebel, S.J.; Johnson, G.P.; Perkins, M.E.; Davis, S.W.; Winslow, J.P.; Paolletti, E.
#journal Virology (1990) 179:247-266
#title The complete DNA sequence of vaccinia virus.
#cross-references MIMD:91021027

CLASSIFICATION
#note annotation: possible protein-coding frames
neither amino acid nor nucleotide sequence is given
#superfamily herpesvirus complement control protein;
complement factor H repeat homology
duplication; extracellular protein

KEYWORDS
FEATURE
20-263
1-19
21-81
86-143
148-201
206-261

#domain signal sequence #status predicted #label SIG
#product C4b-binding protein homolog #status predicted
#label MAR
#domain complement factor H repeat homology #label FH1
#domain complement factor H repeat homology #label FH2
#domain complement factor H repeat homology #label FH3
#domain complement factor H repeat homology #label FH4
#length 263 #molecular-weight 28629 #checksum 8152

SUMMARY

Query Match 19.3%; Score 303; DB 1; Length 263;
Best Local Similarity 34.8%; Pred. No. 1.69e-46;
Matches 65; Conservative 28; Mismatches 76; Indels 18; Gaps 14;

Db 40 ANANNIGDTIELYCLPGRKQKMPYAKCTGTG-TL-ENOCIKRCPSPRODINDQ 96
QY 20 SDQYEPGTQAIYKRCRGYS--LGNVIMCRGKGVNANPLKCKCRKCGHGDPTFT 77
Db 97 LDI-GG-V-FGSSITVSCSGYHLIGESKYSCELTSGTSMWNPAPICSEVCKOSPS 153
QY 78 FTLTGAVVFYGYKAYITCEGYQLLGEIN-YRE-CDTGS-WTNDIPLCEYVKKLPVTA 133
Db 154 ISNGR-HNGYE-DE-YTDSVYVYSCNSGYSLTNSGVYCSGGE-WS-DPFCQVYKRP 207
QY 134 PENKIYVSAMEPDREYHFGQAVFVNCNGYKIEGDEMHCSDDGFWSKPKPCVCEISCK 193
Db 208 HPTISNG 214
QY 194 SPDYING 200

RESULT 11

ENTRY A42755 #type complete
TITLE P-selectin precursor - mouse
ALTERNATE_NAMES CD62; granule membrane protein 140; PADGEM
ORGANISM #formal_name Mus musculus #common_name house mouse
DATE 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 10-Sep-1997

ACCESSIONS A42755; A44899
REFERENCE A42755
#authors Weller, A.; Isenmann, S.; Vestweber, D.
#journal J. Biol. Chem. (1992) 267:15176-15183
#title Cloning of the mouse endothelial selectins. Expression of both E- and P-selectin is inducible by tumor necrosis factor alpha.

#cross-references MIMD:92340571
#accession A42755
#molecule-type mRNA
#status
#residues 1-768 #label WEL
#cross-references GB:M87861; NID:q200552; PID:q200553
#experimental_source endothelial cells
#note Sequence extracted from NCBI backbone (NCBIP:109467)

REFERENCE A44899
#authors Sanders, W.E.; Wilson, R.W.; Ballantyne, C.M.; Beaudet, A.L.
#journal Blood (1992) 80:795-800
#title Molecular cloning and analysis of in vivo expression of murine P-selectin.
#cross-references MIMD:92345617
#accession A44899
#status nucleic acid sequence not shown

#cross-references MIMD:92345617
#accession A44899
#status nucleic acid sequence not shown

##molecule-type mRNA
 ##residues 1-723, 'E', 725-768 ##label SAN
 ##cross-references GB:M72332; NID:9193565; PID:9193566
 ##note Sequence extracted from NCBI Backbone (NCBIP:109900)
 CLASSIFICATION #superfamily complement factor H repeat homology; EGF
 homology
 cell adhesion; glycoprotein; phospholipid; phosphoprotein; transmembrane protein

KEYWORDS

FEATURE
 1-41
 42-768 #domain signal sequence #status predicted #label SIG\
 163-194 #product P-selectin #status predicted #label MAM\
 200-257 #domain EGF homology #label EGF\
 262-319 #domain complement factor H repeat homology #label FH01\
 324-381 #domain complement factor H repeat homology #label FH02\
 386-443 #domain complement factor H repeat homology #label FH03\
 448-505 #domain complement factor H repeat homology #label FH04\
 510-567 #domain complement factor H repeat homology #label FH05\
 580-637 #domain complement factor H repeat homology #label FH06\
 642-699 #domain complement factor H repeat homology #label FH07\
 710-733 #domain transmembrane #status predicted #label TMN\
 734-768 #domain intracellular #status predicted #label INT\
 45,54,107,212,347,
 398,456,467,603,
 654,661,679 #binding-site carbohydrate (Asn) (covalent) #status predicted

SUMMARY #length 768 #molecular-weight 83098 #checksum 2334

Query Match 19.3%; Score 303; DB 2; Length 768;
 Best Local Similarity 33.1%; Pred. No. 1.69e-46;
 Matches 39; Conservative 28; Mismatches 46; Indels 5; Gaps 5;

Db 220 ESNFSCITSCAGYLDGDEL-OCLASGWTNPNPKDAVQCSLEAPPHGTW-AC-M 276
 86 FEIGVAVYTCNEGVLLEINRYECDDTG-WTNDIPICEVVKCLPVAPENKIVSSAM 144
 277 HPIAFAVDSCKFEQCPGYRARGSNLHCTGSGMSEPLPCEALACEPPPIPHGS 334
 145 EPDREYHFQAVRFVNCNSKIKEDDEMHCSDDGFMSKPKVCVEISKSPDV-INGS 201

RESULT 12
 ENTRY 153821 #type complete
 TITLE P-selectin - rat
 ORGANISM #formal_name Rattus norvegicus #common_name Norway rat
 DATE 29-May-1998 #sequence_revision 29-May-1998 #text_change 13-Aug-1999

ACCESSIONS
 REFERENCE 153821
 #authors Auchampach, J.A.; Oliver, M.G.; Anderson, D.C.; Manning, A.M.
 #journal Gene (1994) 145:251-255
 #title Cloning, sequence comparison and in vivo expression of the gene encoding rat P-selectin.
 #cross-references MIMD:94333817

#accession 153821
 ##status Preliminary; translated from GB/EMBL/DBJ
 ##molecule-type mRNA
 ##residues 1-768 #label RES
 #cross-references GB:I23088; NID:9349552; PID:9349553
 CLASSIFICATION #superfamily complement factor H repeat homology; C-type lectin homology; EGF homology

FEATURE
 32-158 #domain C-type lectin homology #label LCN\
 163-194 #domain EGF homology #label EGF\
 262-319 #domain complement factor H repeat homology #label FH1\
 510-567 #domain complement factor H repeat homology #label FH2\
 580-637 #domain complement factor H repeat homology #label FH3\
 642-699 #domain complement factor H repeat homology #label FH4\
 #domain complement factor H repeat homology #label FH5
 #length 768 #molecular-weight 83516 #checksum 4885

SUMMARY
 Query Match 19.3%; Score 303; DB 2; Length 768;
 Best Local Similarity 29.3%; Pred. No. 1.69e-46;
 Matches 49; Conservative 33; Mismatches 75; Indels 10; Gaps 10;

Db 290 FECOPGYRMGRSDIILHCDSSGOW-S-EPLPCEALAC-EPLSEPLHSGMDCFSTGAFGY 346
 32 YKCRGRIYSLGVAVINWCKR-GEWVALNPLRCKQKPCGHDPDTPF-GTFT-LTGSANFEY 88
 347 NSCTFCTEGEVLINGNDIAH-CADLQGTAPAPVCEALQCGEPVPVSKQ-VECS-DPF 403
 89 GYKAVYTCNEGVLLEINRYEC-DTGDWTNDIPICEVVKCLPVAPENKIVSSAMEPD 147
 404 GPLTKQSACSFSCDGSLLVGNASVIRCLATGMSDAPECAVSTP 450
 148 RHYHGAIVRFVNCNSKIKEDDEMHCSDDGFMSKPKVCVEISKSPDV-INGS 194

RESULT 13
 ENTRY NBS
 TITLE #type complete
 ALTERNATE_NAMES apolipoprotein H precursor - mouse
 ORGANISM 50K serum glycoprotein; activated protein C-binding protein;
 DATE #formal_name Mus musculus #common_name house mouse
 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 28-May-1999

ACCESSIONS
 REFERENCE A43286
 #authors Nonaka, M.; Matsuda, Y.; Shirotshi, T.; Moriwaki, K.; Nonaka, M.; Natsume-Sakai, S.
 #journal Genomics (1992) 13:1082-1087
 #title Molecular cloning of mouse beta-2-glycoprotein I and mapping of the gene to chromosome 11.
 #cross-references MIMD:92372000
 #accession A43286

##molecule-type mRNA
 ##residues 1-252, 'A', 254-277, 'N', 279-345 #label NON
 ##cross-references GB:D10056
 ##note the authors translated the codon ACT for residue 253 as Ala and ATG for residue 278 as Asn
 sequence extracted from NCBI Backbone (NCBIP:111791, NCBIP:111794) and corrected to correspond with the nucleotide translation

REFERENCE JC2243
 #authors Sellier, G.C.; Steel, D.M.; Zafiroopoulos, A.; Seery, L.T.; Whitehead, A.S.
 #journal Biochem. Biophys. Res. Commun. (1994) 200:1521-1528
 #title Characterization, expression and evolution of mouse beta2-2-glycoprotein I (apolipoprotein H).
 #cross-references MIMD:94242017
 #accession JC2243

##molecule-type mRNA
 ##residues 1-251, 'R', 253-345 #label SEL
 ##cross-references GB:S70439; NID:9546780; PIDN:AAB30789.1; PID:9546781
 #experimental-source liver

GENETICS
 #gene B2gpl
 #map-position 11
 CLASSIFICATION #superfamily apolipoprotein H; complement factor H repeat homology
 #lipid binding; monomer; plasma; VLDL

FEATURE
 1-19 #domain signal sequence #status predicted #label SIG\
 20-345 #product apolipoprotein H #status predicted #label MAM\
 23-79 #domain complement factor H repeat homology #label FH1\
 84-137 #domain complement factor H repeat homology #label FH2\
 142-200 #domain complement factor H repeat homology #label FH3\
 205-260 #domain complement factor H repeat homology #label FH4\
 264-325 #domain complement factor H repeat homology #label FH5
 23-66, 51-79, 84-124,
 110-137, 142-188,
 174-200, 205-248,
 234-260, 264-315,
 300-307, 325-345,
 105, 117, 162, 183,
 193

#disulfide-bonds #status predicted\
 #binding-site carbohydrate (Asn) (covalent) #status

SUMMARY #length 345 #molecular-weight 38618 #checksum 2016

Query Match 18.9%; Score 297; DB 1; Length 345;
Best Local Similarity 29.0%; Pred. No. 4,556-45;
Matches 51; Conservative 37; Mismatches 78; Indels 10; Gaps 8;

Db 40 SYDPOIYYSCKPGVSGMRFTCLPLTGMW-PINTLR-CVPRVCPAGLLEGIYRY 97
QY 23 TPEGQAIYKCRPGIRSGNVIW-VCR-KGEMVALNPLRKCKRPGCGHGTPTFTL 80

Db 98 TS---FEYKNTISFACNPGF-FLNGTSSCKTEEGKWSPDIPACARITCPPPVKFAILL 153
QY 81 TGGNVFEYVKAVYTCNKGYYOLLGEINRECDTDG-WTNDIPICEVYKCLPYTAPENGKI 139

Db 154 KIRPSAGNNSLYQDTVYKCLPHFAMINDVYMCTEGQNTFRL-PECLEVCKPFP 208
QY 140 VSSAMPEPREYHGOAVRFVCSNGYKIEGDEMHCSDDGFWSKRPCKVEISCKSP 195

RESULT 14
ENTRY A35046 #type complete
TITLE endothelial leukocyte adhesion molecule 1 precursor - human
ORGANISM #formal_name Homo sapiens #common_name man
DATE 27-Jul-1990 #sequence_revision 27-Jul-1990 #text_change 16-Jul-1999

ACCESSIONS
REFERENCE A38615; A35046
#authors Collins, T.; Williams, A.; Johnston, G. I.; Kim, J.; Eddy, R.; Shows, T.; Gimbrone Jr., M. A.; Bevilacqua, M. P.
#journal J. Biol. Chem. (1991) 266:2466-2473
#title Structure and chromosomal location of the gene for endothelial-leukocyte adhesion molecule 1.
#cross-references MIM:91115870
#accession A38615
#status preliminary
#molecule_type DNA
#residues 1-610 #label COL
#cross-references GB:M61893; GB:M58017; NID:g182043; PID:g182046
REFERENCE A35046
#authors Hession, C.; Osborn, L.; Goff, D.; Chi-Rosso, G.; Vassallo, C.; PASEK, M.; Pittack, C.; Tizard, R.; Goetz, S.; McCarthy, K.; Hopple, S.; Lobb, R.
#journal Proc. Natl. Acad. Sci. U.S.A. (1990) 87:1673-1677
#title Endothelial leukocyte adhesion molecule 1: direct expression cloning and functional interactions.
#cross-references MIM:90175359
#accession A35046
#status preliminary
#molecule_type mRNA
#residues 1-610 #label HES
#cross-references GB:M30640; NID:g182047; PID:g182048
CLASSIFICATION #superfamily C-type lectin homology; complement factor H repeat homology
KEYWORDS transmembrane protein
FEATURE
12-138 #domain C-type lectin homology #label LCN\
180-237 #domain complement factor H repeat homology #label FH01\
242-299 #domain complement factor H repeat homology #label FH02\
304-362 #domain complement factor H repeat homology #label FH03\
367-425 #domain complement factor H repeat homology #label FH04\
430-488 #domain complement factor H repeat homology #label FH05\
493-547 #domain complement factor H repeat homology #label FH06
SUMMARY #length 610 #molecular-weight 66655 #checksum 542

Query Match 18.9%; Score 298; DB 2; Length 610;
Best Local Similarity 27.4%; Pred. No. 2,636-45;
Matches 45; Conservative 33; Mismatches 79; Indels 7; Gaps 6;

Db 210 CDRGLPSSMETMOCSSGEMSA--PIPCNVVECDATNPANGFVECFONPGSEFPMNTT 267
QY 34 CRPGYRSLGNVIMVCRK-GEWVALNPLRKCKRPGCGHGTPTFT-LTGNVFEYGVK 91

Db 268 CTFDEEGEFLMAQSL-OCTSSGNWMDNEKPTCKATVCAVAPQONGS-VRCSSHPAGEF 325
QY 92 AVYTCTNEGYYOLLGEINRECDTDG-WTNDIPICEVYKCLPYTAPENGKIVSSAMPEPREY 150

Db 326 TFKSSCNFTCEEGFMLOGPAVECTTQGGWTOQIPVCEAFQCTA 369
QY 151 HFGQAVRFVCSNGIKITGDEMHCSDDGFWSKRPCKVEISCKS 194

Search completed: Thu Jun 8 21:37:54 2000
Job time : 17 secs.

Db 268 CTFDEEGEFLMAQSL-OCTSSGNWMDNEKPTCKATVCAVAPQONGS-VRCSSHPAGEF 325
QY 92 AVYTCTNEGYYOLLGEINRECDTDG-WTNDIPICEVYKCLPYTAPENGKIVSSAMPEPREY 150

Db 326 TFKSSCNFTCEEGFMLOGPAVECTTQGGWTOQIPVCEAFQCTA 369
QY 151 HFGQAVRFVCSNGIKITGDEMHCSDDGFWSKRPCKVEISCKS 194

RESULT 15
ENTRY A32606 #type complete
TITLE endothelial leukocyte adhesion molecule precursor - human
ORGANISM #formal_name Homo sapiens #common_name man
DATE 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 16-Jul-1999

ACCESSIONS
REFERENCE A32606
#authors Bevilacqua, M. P.; Stengelin, S.; Gimbrone Jr., M. A.; Seed, B. Science (1989) 243:1160-1165
#journal Endothelial leukocyte adhesion molecule 1: an inducible receptor for neutrophils related to complement regulatory proteins and lectins.
#title
#cross-references MIM:89162047
#accession A32606
#status preliminary
#molecule_type mRNA
#residues 1-610 #label BEV
#cross-references GB:M24736; NID:g537523; PID:g537524
CLASSIFICATION #superfamily EGF homology; complement factor H repeat homology
KEYWORDS transmembrane protein
FEATURE
1-21 #domain signal sequence #status predicted #label SIG\
22-610 #product endothelial leukocyte adhesion molecule 1
#status predicted #label MAT\
143-174 #domain EGF homology #label EGF\
180-237 #domain complement factor H repeat homology #label FH07\
242-299 #domain complement factor H repeat homology #label FH08\
304-362 #domain complement factor H repeat homology #label FH09\
367-425 #domain complement factor H repeat homology #label FH10\
430-488 #domain complement factor H repeat homology #label FH11\
493-547 #domain complement factor H repeat homology #label FH12
SUMMARY #length 610 #molecular-weight 66681 #checksum 746

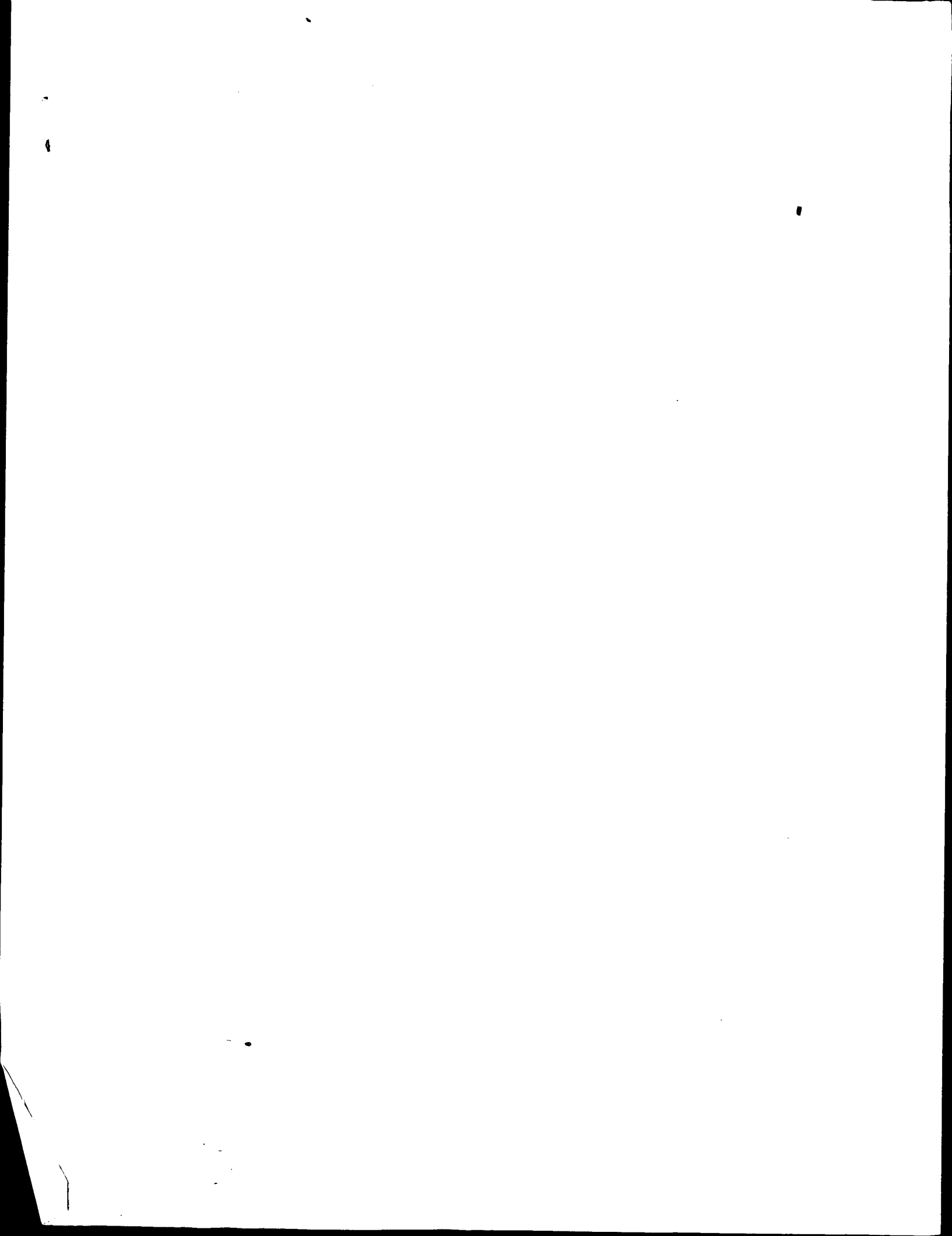
Query Match 18.9%; Score 298; DB 2; Length 610;
Best Local Similarity 27.4%; Pred. No. 2,636-45;
Matches 45; Conservative 33; Mismatches 79; Indels 7; Gaps 6;

Db 210 CDRGLPSSMETMOCSSGEMSA--PIPCNVVECDATNPANGFVECFONPGSEFPMNTT 267
QY 34 CRPGYRSLGNVIMVCRK-GEWVALNPLRKCKRPGCGHGTPTFT-LTGNVFEYGVK 91

Db 268 CTFDEEGEFLMAQSL-OCTSSGNWMDNEKPTCKATVCAVAPQONGS-VRCSSHPAGEF 325
QY 92 AVYTCTNEGYYOLLGEINRECDTDG-WTNDIPICEVYKCLPYTAPENGKIVSSAMPEPREY 150

Db 326 TFKSSCNFTCEEGFMLOGPAVECTTQGGWTOQIPVCEAFQCTA 369
QY 151 HFGQAVRFVCSNGIKITGDEMHCSDDGFWSKRPCKVEISCKS 194

Search completed: Thu Jun 8 21:37:54 2000
Job time : 17 secs.



Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 19 EDCNLEPPRRNELLTGSMSOTVPEGQAIYKCRPGRSIGNVIWCRKKEWALNPLR 78
QY 1 EDCNLEPPRRNELLTGSMSOTVPEGQAIYKCRPGRSIGNVIWCRKKEWALNPLR 60
Db 79 KCQRRPCGHPDTPRGFTLLTGNGVFEYGVAVYTCNNGYOLLGEINRECDTDGWTNDI 138
QY 61 KCQRRPCGHPDTPRGFTLLTGNGVFEYGVAVYTCNNGYOLLGEINRECDTDGWTNDI 120
Db 139 PICEVVKLPYAPENGKIVSSAMEPDRHYHFGQAVFVCSGKIKEGDEMHCSDDGFV 198
QY 121 PICEVVKLPYAPENGKIVSSAMEPDRHYHFGQAVFVCSGKIKEGDEMHCSDDGFV 180
Db 199 SKERKPCVEISCKSPDVINGSPISOKI 225
QY 181 SKERKPCVEISCKSPDVINGSPISOKI 207

RESULT 2
ID Q28085 PRELIMINARY; PRT: 669 AA.

AC Q28085;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-NOV-1999 (TREMblrel. 12, Last annotation update)
DE CCP MODULES 3-12, WITH PARTS OF CCP 2 AND 13 (FRAGMENT).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae;
RN Bovine; Bos.
[1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE: 96202005.
RA SOAMES C.J., DAY A.J., SIM R.B.;
RT "Prediction from sequence comparisons of residues of factor H involved
in the interaction with complement component C3b."
RL Biochem. J. 315:523-531(1996).
DR EMBL: X98697; CA67257.1; -;
DR HSSP: P10998; 1VVC.
PRFAM: PF00084; sush1. 11.
FT NON_TER 1
FT NON_TER 669
SQ SEQUENCE 669 AA: 75683 MW: FAFOD174 CRC32;

Query Match 43.2%; Score 679; DB 6; Length 669;
Best Local Similarity 66.1%; Pred. No. 4.01e-156;
Matches 84; Conservative 21; Mismatches 22; Indels 0; Gaps 0;

Db 7 LAGNDFEYGAQVYTCDEGYOVAGMNPREDCTNGTNDIPICEVVKLPYTERPENGKI 66
QY 80 LAGNDFEYGAQVYTCDEGYOVAGMNPREDCTNGTNDIPICEVVKLPYTERPENGKI 139
Db 67 FSPALDEDTYTGAVYTCNNGYOLLGEINRECDTDGWTNDI 126
QY 140 VSSAMEPDRHYHFGQAVFVCSGKIKEGDEMHCSDDGFVSKERKPCVEISCKSPDVIN 199
Db 127 GOAVLPK 133
QY 200 GSPISOK 206

RESULT 3
ID Q91275 PRELIMINARY; PRT: 1053 AA.

AC Q91275;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-NOV-1999 (TREMblrel. 12, Last annotation update)
DE COMPLEMENT REGULATORY PLASMA PROTEIN.
OS Paratubax nebulifer (barred sand bass).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
OC Neopterygii; Teleostei; Euteleostei; Acanthopterygii; Percomorpha;

OC Perciformes; Percoidae; Serranidae; Paratubax.

RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE: 94318039.
RA DAHMEN A., KAI DOH T., ZIPPEL P.F., GIGLI T.;
RT "Cloning and characterization of a cDNA representing a putative
complement-regulatory plasma protein from barred sand bass (Paratubax
nebulifer)."
RL Biochem. J. 301:391-397(1994).
DR EMBL: U21703; AAA92556.1; -;
DR HSSP: P08603; 1HFN.
PRFAM: PF00084; sush1. 16.
SQ SEQUENCE 1053 AA: 117597 MW: 0D688EDB CRC32;

Query Match 32.0%; Score 503; DB 13; Length 1053;
Best Local Similarity 41.8%; Pred. No. 3.14e-106;
Matches 76; Conservative 29; Mismatches 68; Indels 9; Gaps 7;

Db 47 EASYGGQVRVGVGVY-S-GFEKLVCEGKMETRGA--KCQRRSCGHPEDAQFADPHL 102
QY 21 DQTPDEGQAIYKCRPGRSIGNVIWCRKKEWALNPLRKCQRRPCGHPDTPRGFTLL 80
Db 103 AEGNDFVSGSKVYTCQKGYOVSRINRRCVARGMDGVVPCSSQCC-PATHDNVVOY 161
QY 81 TGNVFEYGVAVYTCNNGYOLLGEINRECDTDGWTNDIPICEVVKLPYAPENGKIV 140
Db 162 ICG--PE-EATFGVNVPRSCSRSEIILDGSELYCDERGMVSGVPCKAITCAIPIEN 218
QY 141 SSAMEPDRHYHFGQAVFVCSGKIKEGDEMHCSDDGFVSKERKPCVEISCKSPDVIN 199
Db 219 GN 220
QY 200 GS 201

RESULT 4
ID Q9WRU2 PRELIMINARY; PRT: 645 AA.

AC Q9WRU2;
DT 01-NOV-1999 (TREMblrel. 12, Created)
DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)
DT 01-NOV-1999 (TREMblrel. 12, Last annotation update)
DE COMPLEMENT BINDING PROTEIN.
OS Macaca mulatta rhadinovirus 17577.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Rhadinovirus.
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE: 99174001.
RA SEARLES R.P., BERGOUAN E.P., AXTHELM M.K., WONG S.W.;
RT "Sequence and genomic analysis of a Rhesus macaque rhadinovirus with
similarity to Kaposi's sarcoma-associated herpesvirus/human
herpesvirus 8."
RL J. Virol. 73:3040-3053(1999).
DR EMBL: AF083501; AAD21332.1; -;
SQ SEQUENCE 645 AA: 71526 MW: 60FB82D6 CRC32;

Query Match 23.1%; Score 364; DB 14; Length 645;
Best Local Similarity 32.6%; Pred. No. 6.97e-68;
Matches 61; Conservative 40; Mismatches 71; Indels 15; Gaps 12;

Db 337 EKYSVGSVELICRPGRTKQSTVSECLSNGTAPNA--KCHRRKCPDPELLNGEYI 394
QY 22 QTYPEGQAIYKCRPGRSIGNVIW-C-RKGEWVALNPLRKCQRRPCGHPDTPRGFT 79
Db 395 VTSGEDAFKYGTNITKCNNGYOLLGSNVRICMKKDDLTVMERAPICDIECKKPPQ 454
QY 80 LTGG-INFEYGVAVYTCNNGYOLLGE-IN-Y--RCC-DTWDITNDIPICEVVKLPYTA 133
Db 455 ITNGKY--HGVK--DF-YOYLDVTTFSCNDRFSLVSGDEMTTCISNT--WNPFPKQIC 509
QY 134 PENGKIVSSAMEPDRHYHFGQAVFVCSGKIKEGDEMHCSDDGFVSKERKPCVEISCK 193

Db	510	APRANG 516
Qy	194	SPDYNG 200
RESULT	5	
ID	040912	PRELIMINARY; PRT; 550 AA.
AC	040912;	
DT	01-JAN-1998 (TREMBlrel. 05, Created)	
DT	01-JAN-1998 (TREMBlrel. 05, Last sequence update)	
DT	01-NOV-1999 (TREMBlrel. 12, Last annotation update)	
DE	ORF 04.	
OS	Kaposi's sarcoma-associated herpesvirus (KSHV) (Human herpesvirus 8).	
OC	viruses; dsDNA viruses, no RNA stage; Herpesviridae;	
OC	Gammapherpesvirinae; Rhadinovirus.	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RX	MEDLINE; 97296220.	
RA	NEIPEL F., ALBRCHT J.C., FLECKENSTEIN B.;	
RT	"Cell-homologous genes in the Kaposi's sarcoma-associated rhadinovirus	
RT	human herpesvirus 8: determinants of its pathogenicity?";	
RL	J. Virol. 71:4187-4192(1997).	
DR	EMBL; D93872; AAB62602.1; -.	
DR	HSSP; P10998; IYVC.	
DR	PFAM; PF00084; sush1; 4.	
SQ	SEQUENCE 550 AA; 60648 MW; 5580989B4 CRC32;	
Query Match	21.9%; Score 344; DB 14; Length 550;	
Best Local Similarity	33.9%; Pred. No. 1,73e-62;	
Matches	61; Conservative 38; Mismatches 64; Indels 17; Gaps 13	
Db	52	RORSGYTTAARNITATCTGGGTW-S-EPTATCNKSKCPNPEIONGKATIFHGQDALKYG 109
Qy	33	KCRPGYRSLG- NVIMWC- RKGEWALNPLRCKQKPCHPDPTFGFTLLGG- NVEYG 89
Db	110	ANISVNCNEGFLVGREYVRCIMGASQGMWSSSPCEKEKC-H-R-P--KIENGDF 163
Qy	90	KVAYTCEGQQLG-E-INTRECTDG---WTNDIPICEYVKCLPYAPENGRIVSSAM 144
Db	164	KPKDYEYNDVAHFECEGTYLVGPHSICAAVNNWTSMNPTCELAGCKFPSTHGYPI 223
Qy	145	EPDREYH-FGQAVRFVNCNGYKIEGDEMHSCSDDGFMSEKPKCWEISCKSPDYNGSPI 203
RESULT	6	
ID	P88903	PRELIMINARY; PRT; 550 AA.
AC	P88903;	
DT	01-MAY-1997 (TREMBlrel. 03, Created)	
DT	01-MAY-1997 (TREMBlrel. 03, Last sequence update)	
DT	01-NOV-1999 (TREMBlrel. 12, Last annotation update)	
DE	ORF 4.	
OS	Kaposi's sarcoma-associated herpesvirus (KSHV) (Human herpesvirus 8).	
OC	viruses; dsDNA viruses, no RNA stage; Herpesviridae;	
OC	Gammapherpesvirinae; Rhadinovirus.	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RX	MEDLINE; 97121480.	
RA	RUSO J.J., BOHEWZY R.A., CHIEN M.C., CHEN J., YAN M., MADDALENA D.,	
RA	PARKY J.P., PERUTZI D., EDLMAN I.S., CHANG Y., MOORE P.S.;	
RT	"Nucleotide sequence of the Kaposi sarcoma-associated herpesvirus	
RT	(HHV8)";	
RL	Proc. Natl. Acad. Sci. U.S.A. 93:14862-14867(1996).	
RN	[2]	
RP	SEQUENCE FROM N.A.	
RX	MEDLINE; 97094384.	
RA	MOORE P.S., BASHOFF C., WEISS R.A., CHANG Y.;	
RT	"Molecular mimicry of human cytokine and cytokine response pathway	
RT	genes by KSHV";	
RL	Science 274:1739-1744(1996).	
DR	EMBL; U75698; AAC57082.1; -.	
DR	HSSP; P10998; IYVC.	
DR	PFAM; PF00084; sush1; 4.	
SQ	SEQUENCE 550 AA; 60687 MW; CE22ACAC CRC32;	

Query Match	21.8%	Score 343;	DB 14;	Length 550;
Best Local Similarity	33.9%	Pred. No. 3,22e-62;		
Matches	61;	Conservative	38;	Mismatches 64; Indels 17; Gaps 13;
Db	52	RCRSGYTYANINATATCLOGGTW-S-EPTATCKKSKCPNPGETIONKGVIFHGGODALRTG	109	
Qy	33	KCRPGYKSLG-NVIMVC-RKGEWALNPRLKCKCRPCGHPDPTFGFTLTGG-NVFEYG	89	
Db	110	ANISYNEGEFLVGRGVYXCMIGASGQAMWSSSPFECKEC-H-R-P---KIKNGDF	163	
Qy	90	VKAVYTCNEGVLQGLG-E-INTRECTDGG---WTINDPICEVWKLPTAPANGKIVSSAM	144	
Db	164	KPDADYEYNAAVFECNEGTYTLVGPISHIACAVANNMTSMPTCELAGKFPSTHGP	223	
Qy	145	EPDREYH-FGGAVAFVNCYKIGEDDMHCSDGGFPMKSEKPKCVELISCKSPDINSPI	203	
RESULT	7	PRELIMINARY;	PRT;	360 AA.
ID	Q9YTO8			
AC	Q9YTO8			
DT	01-MAY-1999 (Tremblrel. 10, Created)			
DT	01-MAY-1999 (Tremblrel. 10, Last sequence update)			
DT	01-NOV-1999 (Tremblrel. 12, Last annotation update)			
DE	COMPLEMENT CONTROL PROTEIN HOMOLOG CCPH.			
OS	Atefline herpesvirus 3.			
OC	Viruses; dsDNA viruses, no RNA stage; Herpesviridae;			
CC	Gammaherpesvirinae; Rhadinovirus.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-73;			
RA	ALBRECHT J.-C., FLECKENSTEIN B.;			
RT	"Primary structure of the Herpesvirus Ateles Genome.";			
RL	Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AF083424; AACG5530.1; -			
DR	HSSP; P10998; IYVC.			
SD	SEQUENCE 360 AA; 40208 MW; F66C7ADA CRC32;			
Query Match	20.3%;	Score 320;	DB 14;	Length 360;
Best Local Similarity	37.6%;	Pred. No. 4.56e-56;		
Matches	71;	Conservative	28;	Mismatches 72; Indels 18; Gaps 13;
Db	40	SSGSYPNGTTLQVTCRKQYIGRQIQIQTVCVGNM-TV-P-NECCRRKCSIPADLLNGHYT	96	
Qy	20	SDQYTPGSTQAIYKCRPGYRSIGNVIMVCRKGEWALNPRLKCKCRPGHPDPTFGFTT	79	
Db	97	VNG-NLY-VGASITTYCTNGYQLGSPF-SSCLLGGDGRVNMWTPRPICETTKCKKPPTI	153	
Qy	80	ILGGNVFEYGAIVATCNEGVLQGLGEIYRRC--DTDG--WTINDPICEVWKLPTAP	134	
Db	154	ANGT--HTNIK-E-YYTLDIAVATYSCNDETRKLITLTPSSKOCSETGRWVPDEETKCEKV	209	
Qy	133	ENGKIVSSAMEPDRHYHGGAVAFVNCYKIE--GDEMHCSDDGFW-SKPKRCAEIS	191	
Db	210	CKIPQVANG 218		
Qy	192	CKSPDIVNG 200		
RESULT	8	PRELIMINARY;	PRT;	259 AA.
ID	P87616			
AC	P87616;			
DT	01-MAY-1997 (Tremblrel. 03, Created)			
DT	01-MAY-1997 (Tremblrel. 03, Last sequence update)			
DT	01-NOV-1999 (Tremblrel. 12, Last annotation update)			
DE	41KP FRAGMENT FROM LEFT END OF GENOME.			
OS	D17L OR C17L.			
OC	Compos virus (CPV).			
CC	Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;			
OC	Orthopoxvirus.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-GRI-90;			

```

RX MEDLINE; 97068532.
RA SAFRONOV P.F., PETROV N.A., RIAZANKINA O.I., TOTMENIN A.V.,
RA SHCHELKHUNOV S.N., SANDAKHCHIV L.S.,
RT "Genes of a circle of hosts for the cowpox virus.";
RL Dokl. Akad. Nauk 349:829-833(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-GRI-90;
RX MEDLINE; 98229462.
RA SHCHELKHUNOV S.N., SAFRONOV P.F., TOTMENIN A.V., PETROV N.A.,
RA RIAZANKINA O.I., GUTOMOV V.V., KOTVAL G.J.;
RT "The genomic sequence analysis of the left and right species-specific
RT terminal region of a cowpox virus strain reveals unique sequences and
RT a cluster of intact ORFs for immunomodulatory and host range
RT proteins.";
RL Virology 243:432-460(1998).
DR EMBL; X94355; CAA64102.1; -.
DR EMBL; Y11842; CAA72567.1; -.
DR HSSP; P10998; IYVC.
DR PFAM; PF00084; sush1; 4.
SQ SEQUENCE 259 AA; 28193 MW; 76531F63 CRC32;

Query Match 20.1%; Score 316; DB 14; Length 259;
Best Local Similarity 34.8%; Pred. No. 5,286-55;
Matches 71; Conservative 30; Mismatches 84; Indels 19; Gaps 15;

DB 20 CCBIPSRITMKRKGTVDSHYNIGDTIEYLCPGYRKOKMGPYAKCTGTGM-TL--FN 75
QY 3 CNELPFRNTEILGSMWSDOTYPEGTQAIKCRGYRS--LGNVIMVCRGEMVYALPFR 60
DB 76 QCIKRCPSRPRIDNGOLDI-GG-V-DFGSSITYSCNSGYHLIGESKSYCELSTGSMW 132
QY 61 KCKRCPCGHGDPFCTFTLGTGNEVEYGVKAVYTCNEGYYQLGELN-YRE-CDTGG--W 116
DB 133 NPAPICSEVCKCSPSISNGR--HNGYE-DF-YTDSVYTYSCNSGYSLIGSGVLCSC 188
QY 117 TNDIPICEVVKCLPYAPENGVKIVSSAMEPDRYHFGQAVFCVNSGKYLEGDEMHCS 176
DB 189 GE-WS-DPPTQIIVKCPHPTISNG 210
QY 177 DGFWSKRPCKVEIISCKSPDIVNG 200

RESULT 9
ID 028290 PRELIMINARY; PRT; 754 AA.
AC 028290;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)
DE CELL ADHESION MOLECULE PRECURSOR (FRAGMENT).
GN GMP140.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Carnivora; Fissipedia; Canidae; Canis.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-SPLEEN;
RA MANNING A.M., SANDERS W.E. JR., KUKIELKA G.L., DORE M.,
RA ROSENBLUM C.L., HAWKINS H.L., MICHAEL L.H., EMTMAN M.L., SMITH C.W.,
RA BEAUBERT A.L., ANDERSON D.C.;
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; M88170; AAA63789.1; -.
DR HSSP; P16109; IFSB.
DR PROSITE; PS00615; G_TYPE_LECTIN_1; 1.
DR PFAM; PF00059; lectin_c; 1.
DR Signal; Cell adhesion; glycoprotein.
FT NON_TER 1
FT SIGNAL 1
FT CHAIN 14 13 POTENTIAL.
SQ SEQUENCE 754 AA; 82303 MW; E9570281 CRC32;

Query Match 19.8%; Score 312; DB 6; Length 754;

```

```

Best Local Similarity 30.5%; Pred. No. 6,086-54;
Matches 51; Conservative 36; Mismatches 69; Indels 11; Gaps 10;

DB 212 SFY-CREGYELNPSKLECLASGTWN-KPPR-CVATQC-PLKTEQSGMNLHSEVF 267
QY 30 AIYKCRGYSISLGNVIMVCRKGEWALNPLKRCQKRPCHGPDTP-FGTFT-LTGGNVF 86
DB 268 QYOSCHFSCEEGFALVGP-EVYQCTASGMMTAAPVCEAVACGPKSPYHSGMDCSP-- 324
QY 87 EYGVKAVYTCNEGYYQLGELINRECDTDG-WTNDIPICEVVKCLPYAPENGVKIVSSAME 145
DB 325 SSAPFYNTSCRHCHGAEFRLBEADLVQCTDLCQWTPAPACALQC 371
QY 146 PDREYHFGQAVRFVNSGKYLEGDEMHCSDDGFWSKRPCKVEIISCKSPDIVNG 192

RESULT 10
ID 022328 PRELIMINARY; PRT; 560 AA.
AC 022328;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)
DE COSMID T07H6.
GN T07H6.5.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabdilia; Rhabdilitida;
OC Rhabdilita; Rhabdilitidae; Rhabdilitae; Peloderinae; Caenorhabdilitis.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE; 94150718.
RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,
RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,
RA CRATON M., DEAR S., DU Z., DUBIN R., FAVELLO A., FULTON L.,
RA GARNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,
RA JONES M., KERSHAW J., KIRSTEN J., LAISTER L., LATREILLE P.,
RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
RA PARSONS J., PERCY C., REIKEN L., ROOPRA A., SANDERS D., SHOWNKEEN R.,
RA MALDON N., SMITH A., SONNHAMMER E., STADEN R., SULLIVAN J.,
RA THIRRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RL Nature 368:32-38(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA GEISEL C.;
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA WATERSTON R.;
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U53344; AAA96225.1; -.
DR HSSP; P10998; IYVC.
DR PFAM; PF00084; sush1; 7.
SQ SEQUENCE 560 AA; 61619 MW; 3A10AF63 CRC32;

Query Match 19.7%; Score 310; DB 5; Length 560;
Best Local Similarity 33.6%; Pred. No. 2,066-53;
Matches 51; Conservative 36; Mismatches 51; Indels 14; Gaps 10;

DB 117 AQMGPPD-LR-CKARACDDPDIDENG---LKEGTFEIPHHVYKSCNPGFTLVGSTS-RQ 170
QY 51 GEWALNPLKRCQKRPCHGPDTPFGFTLLGNGVFEYGVKAVYTCNEGYYQLGELINRE 110
DB 171 CSNGEWTNEPANKATECRSPSPFLHGKVVGSLS-T--YQ-S-VVYVSDGGRVLGQ 224
QY 111 CDTDGS-WTNDIPICEVVKCLPYAPENGVKIVSSAMEPDRYHFGQAVRFVNSGKYLEG 169
DB 225 VQICLAEIGNGNEPCEIRCSVLPPLPNG 256

```

OY 170 EEMHCSDDGFWSEKPKVCYEISCKS-PDYING 200

RESULT 11
ID Q28982; PRELIMINARY; PRT; 482 AA.
AC Q28982;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-NOV-1999 (TREMblrel. 12, Last annotation update)
DE E-SELECTIN.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 97075911.
RA WINKLER H., BROSTJAN C., CSIZMADIA V., NATARAJAN G., ANRATHER J.,
RA BACH F.H.;
RT "The intron-exon structure of the porcine E-selectin-encoding gene,"
RL Gene 176:67-72(1996).
DR EMBL; U37521; AAC48680.1; -.
DR HSSP; P16581; IESL.
DR PROSITE; PS00615; C-TYPE_LECTIN_1; 1.
DR PFAM; PF00008; EGF; 1.
DR PFAM; PF00059; lectin.C; 1.
DR PFAM; PF00084; sushi; 4.
DR PRINTS; PR00343; SELECTIN.
KM Lectin; Glycoprotein.
SQ SEQUENCE 482 AA; 52341 MW; 9452897 CRC32;

Query Match 19.6%; Score 308; DB 6; Length 482;
Best Local Similarity 28.9%; Pred. No. 6, 98e-53;
Matches 52; Conservative 38; Mismatches 82; Indels 8; Gaps 6;

Db 196 QSLPWNTTCAECKEKGFEELIGPEHLCTSSGSDGKP--TCKAVTCPTYHPONGDVSC 253
OY 22 QTPEGTQAIYKCRPGYRSLGNVIMVC-RKGEWALNPLRCKQRCGHPDPTFGFTL 80
Db 254 NHSSIGERAVYSTCHFTCAEFGLOGPAQT-ECTAGQMTQAPVCAKCPAVSQPKNG 312
OY 81 TGNVFEYGVAV--WMCNEGYYLLGEINRECDTGD-WTNDIPICEVYKCLPYTAPENG 137
Db 313 -LVFTHSPTGEFTYKSCAFSCFELRGSAQLACTSOGQMTQOEVPSQVYSCSSLEV 371
OY 138 KIVSAMPEDEHYHFGQAVRVCNSGYRIEGDEHMCSDDFGWSKPKVCYEISCKSPDV 197

RESULT 12
ID Q14006; PRELIMINARY; PRT; 657 AA.
AC Q14006;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-NOV-1999 (TREMblrel. 12, Last annotation update)
DE COMPLEMENT H FACTOR (FRAGMENT).
GN HF.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Homiidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 88025472.
RA DAY A.J., RIPOCHE J., LYONS A., MCINTOSH B., HARRIS T.J., SIM R.B.;
RT "Sequence analysis of a cDNA clone encoding the C-terminal end of
RT human complement factor H,"
RL Biosci. Rep. 7:201-207(1987).
DR EMBL; M17517; AAA52016.1; -.
DR HSSP; P08603; 1HFH.
DR PFAM; PF00084; sushi; 11.
FT NON_TER 1
SQ SEQUENCE 657 AA; 74247 MW; F4AB5238 CRC32;

Query Match 19.3%; Score 304; DB 4; Length 657;
Best Local Similarity 27.1%; Pred. No. 7, 96e-52;

Matches 56; Conservative 51; Mismatches 86; Indels 14; Gaps 14;

Db 296 CSQ-PPQIEHGTINSSRSQSEYAHGKLSYNGEGFRISSEHETTCYMGK-W-S-SP-PQ 351
OY 3 CNELPFRNTEILTGSWSDQ-TYPEGTQAIYKCRPGYRSLGNVIMVC-RKGEWALNPLRCK 61
Db 352 CEGLPCKSPPEISHGVVAHMSD-YOYGEVYKCFEGGIDPAIAK-CLGEKWSHP-P 408
OY 62 CQKPCGHPDPTFGFTLNGVNFEEGVAVYTCNMGYOLGELINRECDTGDWTNDIP 121
Db 409 SCIKTDLSLPSEENA-IPNGEKK-D-VIKAGEQVYTYATYIKMGASNVTCINSR-WT 464
OY 122 ICEVYKCLPYTAPENGKIVSSAMEPDREYHFGQAVRVCNSGYRIEGDEHMCSDDFGWS 181
Db 465 -GRPTCDTSCVNPPTYQNAIVSRQ 490
OY 182 KKKPKVCYEISCKSPDYINGSPISQKI 207

Query Match 18.9%; Score 297; DB 4; Length 740;
Best Local Similarity 28.8%; Pred. No. 5, 58e-50;
Matches 51; Conservative 39; Mismatches 75; Indels 12; Gaps 11;

Db 281 AFAYGSSCKFEQOPGYRVRGLMDLRCIDSGHMSA--PLEPTCAISC-EPLSEPVHGSMD 337
OY 23 TYPEGTQAIYKCRPGYRSLGNVIMVC-RKGEWALNPLRCKQRCGHPDPTFGFTL 80
Db 338 SPSLRAFOYDINCSPRCACGFMRLGADIVR-CDNLGQMTAPAPVCALOCODLPVFNAR 396
OY 81 TGG-VNFEYGVAVYTCNMGYOLGELINRECDTGD-WTNDIPICEVYKCLPYTAPENG 138
Db 397 -VNCS-HRPGARRY-QSVCSFTCNGLLIVGASVLOCATGAMNSVSPCECAIPCTP 450
OY 139 IVSSAMEPDREYHFGQAV-RVCNSGYRIEGDEHMCSDDFGWSKPKVCYEISCKS 194

RESULT 14
ID P78361; PRELIMINARY; PRT; 347 AA.
AC P78361;
DT 01-MAY-1997 (TREMblrel. 03, Created)
DT 01-MAY-1997 (TREMblrel. 03, Last sequence update)
DT 01-NOV-1999 (TREMblrel. 12, Last annotation update)
DE DECAY-ACCELERATION FACTOR (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Homiidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE-BRAIN HIPPOCAMPUS;
RA KUMAR V.B., HYUNG C., NAKRA R., WALTERS M., SASSER T., BERNARDO A.;
RL Submitted (FEB-1997) to the EMBL/Genbank/DBJ databases.

 WIRE (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

MSrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Thu Jun 8 21:38:12 2000; MasPar time 8.92 Seconds
 706.532 Million cell updates/sec

Tabular output not generated.

Title: >US-09-316-163-9
 Description: (1-207) from US09316163.pep
 Perfect Score: 1573
 Sequence: 1 EDCNELEPPRRNTETILTSWS.....VEISCKSPDVINGSPISOKI 207

Scoring table: PAM 150
 Gap 11

Searched: 83857 seqs, 30454973 residues

Post-Processing: Minimum Match 0
 Listing first 45 summaries

Database: swiss-prot38
 1:swissprot

Statistics: Mean 42.607; Variance 59.776; scale 0.713

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	1573	100.0	1231	1	CFAB_HUMAN COMPLEMENT FACTOR H PR	0.00e+00
2	1154	73.4	1234	1	CFAB_MOUSE COMPLEMENT FACTOR H PR	6.87e-299
3	1058	19.6	484	1	LEM2_PIG E-SELECTIN PRECURSOR (8.32e-54
4	308	19.6	830	1	LEM3_HUMAN P-SELECTIN PRECURSOR (1.82e-52
5	303	19.3	263	1	VCP_VACCV P-SELECTIN PRECURSOR (1.82e-52
6	303	19.3	768	1	LEM3_MOUSE P-SELECTIN PRECURSOR (1.82e-52
7	303	19.3	768	1	LEM3_RAT P-SELECTIN PRECURSOR (1.82e-52
8	297	18.9	345	1	APDH_MOUSE BETA-2-GLYCOPROTEIN I	7.31e-51
9	298	18.9	610	1	LEM2_HUMAN E-SELECTIN PRECURSOR (3.95e-51
10	292	18.6	345	1	APDH_BOVIN BETA-2-GLYCOPROTEIN I	1.57e-49
11	293	18.6	381	1	DAF_HUMAN COMPLEMENT DECAY-ACCEL	5.35e-49
12	290	18.4	597	1	C4BP_HUMAN C4B-BINDING PROTEIN AL	5.35e-49
13	288	18.3	360	1	CCPH_HSVSA COMPLEMENT CONTROL PRO	1.82e-48
14	288	18.3	612	1	LEM2_MOUSE E-SELECTIN PRECURSOR (1.82e-48
15	286	18.2	549	1	LEM2_MOUSE E-SELECTIN PRECURSOR (6.17e-48
16	287	18.2	551	1	LEM2_RAT E-SELECTIN PRECURSOR (3.35e-48
17	284	18.1	340	1	DAF_PONPY COMPLEMENT DECAY-ACCEL	2.09e-47
18	281	17.9	377	1	MCB_HUMAN MEMBRANE COFACTOR PROT	1.30e-46
19	280	17.8	507	1	DAF_CAVPO COMPLEMENT DECAY-ACCEL	2.39e-46
20	279	17.7	611	1	LEM2_CANFA E-SELECTIN PRECURSOR (4.40e-46
21	276	17.5	1019	1	LFC_TACTR LIMOUS CLOTTING FACTO	2.72e-45
22	272	17.3	958	1	HIG_PROMO LOCOCOMOTION-RELATED PRO	3.07e-44
23	270	17.2	345	1	APDH_HUMAN BETA-2-GLYCOPROTEIN I	1.03e-43

24	270	17.2	610	1	C4BP_BOVIN	C4B-BINDING PROTEIN AL	1.03e-43
25	270	17.2	1033	1	CR2_HUMAN	COMPLEMENT RECEPTOR TY	1.03e-43
26	267	17.0	2039	1	CR1_HUMAN	COMPLEMENT RECEPTOR TY	6.31e-43
27	266	16.9	485	1	LEM2_BOVIN	E-SELECTIN PRECURSOR (1.15e-42
28	264	16.8	345	1	APDH_CANFA	BETA-2-GLYCOPROTEIN I	3.85e-42
29	265	16.8	469	1	C4BP_MOUSE	C4B-BINDING PROTEIN PR	2.11e-42
30	260	16.5	558	1	C4BP_RAT	C4B-BINDING PROTEIN AL	4.27e-41
31	257	16.3	1025	1	CR2_MOUSE	COMPLEMENT RECEPTOR TY	2.58e-40
32	254	16.1	390	1	DAF1_MOUSE	COMPLEMENT DECAY-ACCEL	1.56e-39
33	251	16.0	769	1	LEM3_SHEEP	P-SELECTIN PRECURSOR (9.33e-39
34	238	15.1	407	1	DAF2_MOUSE	COMPLEMENT DECAY-ACCEL	2.09e-35
35	236	15.0	297	1	APDH_RAT	BETA-2-GLYCOPROTEIN I	6.80e-35
36	233	14.8	646	1	LEM3_BOVIN	P-SELECTIN PRECURSOR (3.98e-34
37	232	14.7	372	1	LEM3_RAT	L-SELECTIN PRECURSOR (7.15e-34
38	231	14.7	372	1	LEM1_MOUSE	L-SELECTIN PRECURSOR (1.29e-33
39	220	14.0	372	1	LEM1_PANTR	L-SELECTIN PRECURSOR (7.89e-31
40	220	14.0	372	1	LEM1_MACMO	L-SELECTIN PRECURSOR (7.89e-31
41	220	14.0	372	1	LEM1_PAPHA	L-SELECTIN PRECURSOR (7.89e-31
42	220	14.0	372	1	LEM1_HUMAN	L-SELECTIN PRECURSOR (7.89e-31
43	218	13.9	372	1	LEM1_PONPY	L-SELECTIN PRECURSOR (2.51e-30
44	204	13.0	661	1	F13B_HUMAN	COAGULATION FACTOR XII	7.84e-27
45	197	12.5	330	1	CFH1_HUMAN	COMPLEMENT FACTOR H-LI	4.17e-25

ALIGNMENTS

RESULT ID	1	STANDARD	PRT	1231 AA.
AC	P08603:			
DT	01-AUG-1988 (Rel. 08, Created)			
DT	01-JAN-1990 (Rel. 13, Last sequence update)			
DT	15-FEB-2000 (Rel. 39, Last annotation update)			
DE	COMPLEMENT FACTOR H PRECURSOR.			
GN	HEF OR HE OR CFH.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia;			
OC	Euthera; Primates; Catarrhini; Homiidae; Homo.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=LIVER;			
RA	MEDLINE: 88134059.			
RA	Ripoche J., Day A.J., Harris T.J.R., Sim R.B.;			
RT	"the complete amino acid sequence of human complement factor H.";			
RL	Biochem. J. 249:593-602(1988).			
RN	[2]			
RP	SEQUENCE OF 53-445 FROM N.A.			
RC	MEDLINE: 87054207.			
RA	Schulz T.F., Schwaible W., Stanley K.K., Weiss E., Dierich M.P.;			
RT	"Human complement factor H: isolation of cDNA clones and partial cDNA			
RT	sequence of the 38-kDa tryptic fragment containing the binding site			
RT	for C3b.";			
RT	Eur. J. Immunol. 16:1351-1355(1986).			
RN	[3]			
RP	SEQUENCE OF 226-445 FROM N.A. AND PARTIAL SEQUENCE.			
RC	MEDLINE: 86169701.			
RA	Kristensen T., Wetzel R.A., Tack B.F.;			
RT	"Structural analysis of human complement protein H: homology with C4b			
RT	binding protein, beta-2-glycoprotein I, and the Ba fragment of B2.";			
RL	J. Immunol. 136:3407-3411(1986).			
RN	[4]			
RP	SEQUENCE OF 1047-1231 FROM N.A.			
RC	MEDLINE: 91201892.			
RA	Estaller C., Koistinen V., Schwaible W., Dierich M.P., Weiss E.H.;			
RT	"Cloning of the 1.4-kb mRNA species of human complement factor H			
RT	reveals a novel member of the short consensus repeat family related			
RT	to the carboxy terminal of the classical 150-kDa molecule.";			
RL	J. Immunol. 146:3190-3196(1991).			
RN	[5]			
RP	SEQUENCE OF 19-35.			
RC	MEDLINE: 83048213.			
RA	Sam R.B., Discipio R.G.;			
RT	"Purification and structural studies on the complement-system control			
RT	protein beta 1H (Factor H).";			

```

RN RL Biochem. J. 205:285-293(1982).
RR RP STRUCTURE BY NMR OF 927-985 (SUSHI 16).
RA MEDLINE; 91278097.
RX Norman D.G., Barlow P.N., Baron M., Day A.J., Sim B., Campbell I.D.;
RT "three-dimensional structure of a complement control protein module
in solution".
RL J. Mol. Biol. 219:717-725(1991).
[7]
RR RP STRUCTURE BY NMR OF 264-322 (SUSHI 5).
RA MEDLINE; 92232649.
RX Barlow P.N., Norman D.G., Steinkasserer A., Horne T.J., Pearce J.,
RT Driscoll P.C., Sim B., Campbell I.D.;
RT "Solution structure of the fifth repeat of factor H: a second example
of the complement control protein module."
RL Biochemistry 31:3626-3634(1992).
[8]
RR RP STRUCTURE BY NMR OF 866-985 (SUSHIS 15 AND 16).
RA MEDLINE; 93323119.
RX Barlow P.N., Steinkasserer A., Norman D.G., Kieffer B., Wiles A.P.,
RT "Solution structure of a pair of complement modules by nuclear
magnetic resonance."
RL J. Mol. Biol. 232:268-284(1993).
CC -! FUNCTION: FACTOR H FUNCTIONS AS A COFACTOR IN THE INACTIVATION OF
CC C3B BY FACTOR I AND ALSO INCREASES THE RATE OF DISSOCIATION OF THE
CC C3BB COMPLEX (C3 CONVERTASE) AND THE (C3)NBB COMPLEX (C5
CC CONVERTASE) IN THE ALTERNATIVE COMPLEMENT PATHWAY.
CC -! SIMILARITY: CONTAINS 20 SUSHI (SCR) REPEATS.
CC -! CAUTION: REP.2 SEQUENCE DIFFERS FROM THAT SHOWN FROM POSITION
CC 341 ONWARD DUE TO A FRAMESHIFT.
-----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-slb.ch/announce/
CC or send an email to license@isb-slb.ch).
-----
DR EMBL; Y00716; CAA68704.1; -.
DR EMBL; X04697; CAB41739.1; ALT_FRAME.
DR EMBL; M65294; AAA35946.1; -.
DR PIR; S00254; NBHDH.
DR PIR; S00254; S00254.
DR PIR; S03013; S03013.
DR PDB; IHCC; 15-APR-92.
DR PDB; IHFH; 15-JUL-93.
DR PDB; IHFI; 15-JUL-93.
DR MIM; 134370; -.
DR PFAM; PF00084; sushi. 20.
KW Complement alternate pathway; Plasma; glycoprotein; Repeat; Sushi;
KW Signal; 3d-structure; Polymorphism.
FT SIGNAL 1 18
FT CHAIN 19 1231
FT DOMAIN 20 1229
FT REPEAT 20 81
FT REPEAT 84 142
FT REPEAT 145 206
FT REPEAT 209 263
FT REPEAT 266 321
FT REPEAT 324 386
FT REPEAT 388 443
FT REPEAT 447 506
FT REPEAT 508 565
FT REPEAT 568 624
FT REPEAT 629 685
FT REPEAT 690 745
FT REPEAT 752 804
FT REPEAT 810 865
FT REPEAT 869 927
FT REPEAT 930 985
FT REPEAT 1044 1044
FT SIGNAL 1 18
FT CHAIN 19 1231
FT DOMAIN 20 1229
FT REPEAT 20 81
FT REPEAT 84 142
FT REPEAT 145 206
FT REPEAT 209 263
FT REPEAT 266 321
FT REPEAT 324 386
FT REPEAT 388 443
FT REPEAT 447 506
FT REPEAT 508 565
FT REPEAT 568 624
FT REPEAT 629 685
FT REPEAT 690 745
FT REPEAT 752 804
FT REPEAT 810 865
FT REPEAT 869 927
FT REPEAT 930 985
FT REPEAT 1044 1044
FT SIGNAL 1 18
FT CHAIN 19 1231
FT DOMAIN 20 1229
FT REPEAT 20 81
FT REPEAT 84 142
FT REPEAT 145 206
FT REPEAT 209 263
FT REPEAT 266 321
FT REPEAT 324 386
FT REPEAT 388 443
FT REPEAT 447 506
FT REPEAT 508 565
FT REPEAT 568 624
FT REPEAT 629 685
FT REPEAT 690 745
FT REPEAT 752 804
FT REPEAT 810 865
FT REPEAT 869 927
FT REPEAT 930 985
FT REPEAT 1044 1044
FT SIGNAL 1 18
FT CHAIN 19 1231
FT DOMAIN 20 1229
FT REPEAT 20 81
FT REPEAT 84 142
FT REPEAT 145 206
FT REPEAT 209 263
FT REPEAT 266 321
FT REPEAT 324 386
FT REPEAT 388 443
FT REPEAT 447 506
FT REPEAT 508 565
FT REPEAT 568 624
FT REPEAT 629 685
FT REPEAT 690 745
FT REPEAT 752 804
FT REPEAT 810 865
FT REPEAT 869 927
FT REPEAT 930 985
FT REPEAT 1044 1044
FT SIGNAL 1 18
FT CHAIN 19 1231
FT DOMAIN 20 1229
FT REPEAT 20 81
FT REPEAT 84 142
FT REPEAT 145 206
FT REPEAT 209 263
FT REPEAT 266 321
FT REPEAT 324 386
FT REPEAT 388 443
FT REPEAT 447 506
FT REPEAT 508 565
FT REPEAT 568 624
FT REPEAT 629 685
FT REPEAT 690 745
FT REPEAT 752 804
FT REPEAT 810 865
FT REPEAT 869 927
FT REPEAT 930 985
FT REPEAT 1044 1044
FT SIGNAL 1 18
FT CHAIN 19 1231
FT DOMAIN 20 1229
FT REPEAT 20 81
FT REPEAT 84 142
FT REPEAT 145 206
FT REPEAT 209 263
FT REPEAT 266 321
FT REPEAT 324 386
FT REPEAT 388 443
FT REPEAT 447 506
FT REPEAT 508 565
FT REPEAT 568 624
FT REPEAT 629 685
FT REPEAT 690 745
FT REPEAT 752 804
FT REPEAT 810 865
FT REPEAT 869 927
FT REPEAT 930 985
FT REPEAT 1044 1044
FT SIGNAL 1 18
FT CHAIN 19 1231
FT DOMAIN 20 1229
FT REPEAT 20 81
FT REPEAT 84 142
FT REPEAT 145 206
FT REPEAT 209 263
FT REPEAT 266 321
FT REPEAT 324 386
FT REPEAT 388 443
FT REPEAT 447 506
FT REPEAT 508 565
FT REPEAT 568 624
FT REPEAT 629 685
FT REPEAT 690 745
FT REPEAT 752 804
FT REPEAT 810 865
FT REPEAT 869 927
FT REPEAT 930 985
FT REPEAT 1044 1044
FT SIGNAL 1 18
FT CHAIN 19 1231
FT DOMAIN 20 1229
FT REPEAT 20 81
FT REPEAT 84 142
FT REPEAT 145 206
FT REPEAT 209 263
FT REPEAT 266 321
FT REPEAT 324 386
FT REPEAT 388 443
FT REPEAT 447 506
FT REPEAT 508 565
FT REPEAT 568 624
FT REPEAT 629 685
FT REPEAT 690 745
FT REPEAT 752 804
FT REPEAT 810 865
FT REPEAT 869 927
FT REPEAT 930 985
FT REPEAT 1044 1044
FT SIGNAL 1 18
FT CHAIN 19 1231
FT DOMAIN 20 1229
FT REPEAT 20 81
FT REPEAT 84 142
FT REPEAT 145 206
FT REPEAT 209 263
FT REPEAT 266 321
FT REPEAT 324 386
FT REPEAT 388 443
FT REPEAT 447 506
FT REPEAT 508 565
FT REPEAT 568 624
FT REPEAT 629 685
FT REPEAT 690 745
FT REPEAT 752 804
FT REPEAT 810 865
FT REPEAT 869 927
FT REPEAT 930 985
FT REPEAT 1044 1044
FT SIGNAL 1 18
FT CHAIN 19 1231
FT DOMAIN 20 1229
FT REPEAT 20 81
FT REPEAT 84 142
FT REPEAT 145 206
FT REPEAT 209 263
FT REPEAT 266 321
FT REPEAT 324 386
FT REPEAT 388 443
FT REPEAT 447 506
FT REPEAT 508 565
FT REPEAT 568 624
FT REPEAT 629 685
FT REPEAT 690 745
FT REPEAT 752 804
FT REPEAT 810 865
FT REPEAT 869 927
FT REPEAT 930 985
FT REPEAT 1044 1044
FT SIGNAL 1 18
FT CHAIN 19 1231
FT DOMAIN 20 1229
FT REPEAT 20 81
FT REPEAT 84 142
FT REPEAT 145 206
FT REPEAT 209 263
FT REPEAT 266 321
FT REPEAT 324 386
FT REPEAT 388 443
FT REPEAT 447 506
FT REPEAT 508 565
FT REPEAT 568 624
FT REPEAT 629 685
FT REPEAT 690 745
FT REPEAT 752 804
FT REPEAT 810 865
FT REPEAT 869 927
FT REPEAT 930 985
FT REPEAT 1044 1044
FT SIGNAL 1 18
FT CHAIN 19 1231
FT DOMAIN 20 1229
FT REPEAT 20 81
FT REPEAT 84 142
FT REPEAT 145 206
FT REPEAT 209 263
FT REPEAT 266 321
FT REPEAT 324 386
FT REPEAT 388 443
FT REPEAT 447 506
FT REPEAT 508 565
FT REPEAT 568 624
FT REPEAT 629 685
FT REPEAT 690 745
FT REPEAT 752 804
FT REPEAT 810 865
FT REPEAT 869 927
FT REPEAT 930 985
FT REPEAT 1044 1044
FT SIGNAL 1 18
FT CHAIN 19 1231
FT DOMAIN 20 1229
FT REPEAT 20 81
FT REPEAT 84 142
FT REPEAT 145 206
FT REPEAT 209 263
FT REPEAT 266 321
FT REPEAT 324 386
FT REPEAT 388 443
FT REPEAT 447 506
FT REPEAT 508 565
FT REPEAT 568 624
FT REPEAT 629 685
FT REPEAT 690 745
FT REPEAT 752 804
FT REPEAT 810 865
FT REPEAT 869 927
FT REPEAT 930 985
FT REPEAT 1044 1044
FT SIGNAL 1 18
FT CHAIN 19 1231
FT DOMAIN 20 1229
FT REPEAT 20 81
FT REPEAT 84 142
FT REPEAT 145 206
FT REPEAT 209 263
FT REPEAT 266 321
FT REPEAT 324 386
FT REPEAT 388 443
FT REPEAT 447 506
FT REPEAT 508 565
FT REPEAT 568 624
FT REPEAT 629 685
FT REPEAT 6
```

FT	REPEAT	1047	1103	SUSHI 18.
FT	REPEAT	1108	1164	SUSHI 19.
FT	REPEAT	1166	1229	SUSHI 20.
FT	DISULFID	21	86	BY SIMILARITY.
FT	DISULFID	52	60	BY SIMILARITY.
FT	DISULFID	85	129	BY SIMILARITY.
FT	DISULFID	114	141	BY SIMILARITY.
FT	DISULFID	146	192	BY SIMILARITY.
FT	DISULFID	178	205	BY SIMILARITY.
FT	DISULFID	210	251	BY SIMILARITY.
FT	DISULFID	237	262	BY SIMILARITY.
FT	DISULFID	267	309	BY SIMILARITY.
FT	DISULFID	294	320	BY SIMILARITY.
FT	DISULFID	325	374	BY SIMILARITY.
FT	DISULFID	357	385	BY SIMILARITY.
FT	DISULFID	389	431	BY SIMILARITY.
FT	DISULFID	416	442	BY SIMILARITY.
FT	DISULFID	448	494	BY SIMILARITY.
FT	DISULFID	477	505	BY SIMILARITY.
FT	DISULFID	509	553	BY SIMILARITY.
FT	DISULFID	536	564	BY SIMILARITY.
FT	DISULFID	569	611	BY SIMILARITY.
FT	DISULFID	597	623	BY SIMILARITY.
FT	DISULFID	630	673	BY SIMILARITY.
FT	DISULFID	659	684	BY SIMILARITY.
FT	DISULFID	691	733	BY SIMILARITY.
FT	DISULFID	719	744	BY SIMILARITY.
FT	DISULFID	753	792	BY SIMILARITY.
FT	DISULFID	781	803	BY SIMILARITY.
FT	DISULFID	811	853	BY SIMILARITY.
FT	DISULFID	839	864	BY SIMILARITY.
FT	DISULFID	870	915	BY SIMILARITY.
FT	DISULFID	901	926	BY SIMILARITY.
FT	DISULFID	931	973	BY SIMILARITY.
FT	DISULFID	959	984	BY SIMILARITY.
FT	DISULFID	989	1032	BY SIMILARITY.
FT	DISULFID	1018	1043	BY SIMILARITY.
FT	DISULFID	1048	1091	BY SIMILARITY.
FT	DISULFID	1077	1102	BY SIMILARITY.
FT	DISULFID	1109	1152	BY SIMILARITY.
FT	DISULFID	1138	1163	BY SIMILARITY.
FT	DISULFID	1167	1218	BY SIMILARITY.
FT	DISULFID	1201	1228	BY SIMILARITY.
FT	CARBOHYD	529	559	POTENTIAL.
FT	CARBOHYD	718	718	POTENTIAL.
FT	CARBOHYD	802	802	POTENTIAL.
FT	CARBOHYD	822	822	POTENTIAL.
FT	CARBOHYD	882	882	POTENTIAL.
FT	CARBOHYD	911	911	POTENTIAL.
FT	CARBOHYD	1029	1029	POTENTIAL.
FT	CARBOHYD	1095	1095	POTENTIAL.
FT	CARBOHYD	1201	1228	POTENTIAL.
FT	CARBOHYD	529	559	POTENTIAL.
FT	CARBOHYD	718	718	POTENTIAL.
FT	CARBOHYD	802	802	POTENTIAL.
FT	CARBOHYD	822	822	POTENTIAL.
FT	CARBOHYD	882	882	POTENTIAL.
FT	CARBOHYD	911	911	POTENTIAL.
FT	CARBOHYD	1029	1029	POTENTIAL.
FT	CARBOHYD	1095	1095	POTENTIAL.
FT	CARBOHYD	1201	1228	POTENTIAL.
FT	CARBOHYD	529	559	POTENTIAL.
FT	CARBOHYD	718	718	POTENTIAL.
FT	CARBOHYD	802	802	POTENTIAL.
FT	CARBOHYD	822	822	POTENTIAL.
FT	CARBOHYD	882	882	POTENTIAL.
FT	CARBOHYD	911	911	POTENTIAL.
FT	CARBOHYD	1029	1029	POTENTIAL.
FT	CARBOHYD	1095	1095	POTENTIAL.
FT	CARBOHYD	1201	1228	POTENTIAL.
FT	CARBOHYD	529	559	POTENTIAL.
FT	CARBOHYD	718	718	POTENTIAL.
FT	CARBOHYD	802	802	POTENTIAL.
FT	CARBOHYD	822	822	POTENTIAL.
FT	CARBOHYD	882	882	POTENTIAL.
FT	CARBOHYD	911	911	POTENTIAL.
FT	CARBOHYD	1029	1029	POTENTIAL.
FT	CARBOHYD	1095	1095	POTENTIAL.
FT	CARBOHYD	1201	1228	POTENTIAL.
FT	CARBOHYD	529	559	POTENTIAL.
FT	CARBOHYD	718	718	POTENTIAL.
FT	CARBOHYD	802	802	POTENTIAL.
FT	CARBOHYD	822	822	POTENTIAL.
FT	CARBOHYD	882	882	POTENTIAL.
FT	CARBOHYD	911	911	POTENTIAL.
FT	CARBOHYD	1029	1029	POTENTIAL.
FT	CARBOHYD	1095	1095	POTENTIAL.
FT	CARBOHYD	1201	1228	POTENTIAL.
FT	CARBOHYD	529	559	POTENTIAL.
FT	CARBOHYD	718	718	POTENTIAL.
FT	CARBOHYD	802	802	POTENTIAL.
FT	CARBOHYD	822	822	POTENTIAL.
FT	CARBOHYD	882	882	POTENTIAL.
FT	CARBOHYD			

FT	CARBOHYD	1030	1030	POTENTIAL.
FT	CARBOHYD	1061	1061	POTENTIAL.
FT	CARBOHYD	1225	1225	POTENTIAL.
SEQ	SEQUENCE	1234 AA;	139082 MM;	C5AC02F341B957F7 CRC64;
Query Match				
Best Local Similarity		73.4%;	Score 1154;	DB 1; Length 1234;
Matches 140;		Conservative 33;	Mismatches 33;	Indels 0; Gaps 0;
Db	19	EDCKRPPRENEELLISGMSSEDIYPEGTQATYKCRPGYRTLTIVYCKANGKVASNPSPR	78	
Qy	1	EDCNLEPPRRNEELLTGSMSDQTEGDAIYKCRPGYRSLGNVIMVCRKGEVVALNPLR	60	
Db	79	ICRRKPCGHPGDTPEGSEFLAVGSGFEFGAKVYVTCDDGTYOLLGELDYRCCADGINDI	138	
Qy	61	KCKRPGCGHPGDTPEGTFITLTGNGVNFVEYKAVYTCNCTYLLGELNRECDTMDTNDI	120	
Db	139	PLCEVVKCLPTELENGRIYVSGAELTDDEYFYGVYRRCNSGFKIEGKHEICSENGIM	198	
Qy	121	PICEVVKCLPVTAPENGRIVSSAMPDEYHFGAVRFVCGNSGYKIEGDEMHKCDGDFW	180	
Db	199	SNKPRCVIEILCTPRVENGDDINK	224	
Qy	181	SKERKCVIEISCKSPDIVNGSPISOK	206	
RESULT 3				
ID	LEN2_PIG	STANDARD;	PRT;	484 AA.
AC	P98110;			
DT	01-FEB-1996	(Rel. 33, Created)		
DT	01-FEB-1996	(Rel. 33, Last sequence update)		
DT	15-JUL-1998	(Rel. 36, Last annotation update)		
DE	E-SELECTIN PRECURSOR (ENDOTHELIAL LEUKOCYTE ADHESION MOLECULE 1)			
DE	(ELAM-1) (LEUKOCYTE-ENDOTHELIAL CELL ADHESION MOLECULE 2) (LECAM2)			
DE	(CD62E).			
OS	SELE.			
GN	Sus scrofa (Pig).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;			
OC	Eutheria; Cetartiodactyla; Suina; Suidae; Sus.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE-AORTIC ENDOTHELIUM;			
RA	MEDLINE; 95071392.			
RA	Rollins S.A., Evans M.J., Johnson K.K., Elliot E.A., Squinto S.P.,			
RA	Matlis L.A., Rother R.P.;			
RT	"Molecular and functional analysis of porcine E-selectin reveals a			
RT	potential role in xenograft rejection."			
RL	Biochem. Biophys. Res. Commun. 204:763-771(1994).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE-AORTIC ENDOTHELIUM;			
RA	MEDLINE; 94271236.			
RA	Tsang Y.T.M., Haskard D.O., Robinson M.K.;			
RT	"Cloning and expression kinetics of porcine vascular cell adhesion			
RT	molecule."			
RL	Biochem. Biophys. Res. Commun. 201:805-805(1994).			
CC	-1- FUNCTION: EXPRESSED ON CYTOKINE INDUCED ENDOTHELIAL CELLS AND			
CC	MEDIATES THEIR BINDING TO LEUKOCYTES. THE LIGAND RECOGNIZED BY			
CC	ELAM-1 IS SIALLYL-LEWIS X (ALPHA(1->3)FUCOSYLATED DERIVATIVES OF			
CC	POLYGLYCOSAMINE THAT ARE FOUND AT THE NONREDUCING TERMINI OF			
CC	GLYCOPROTEINS).			
CC	-1- FUNCTION: PLAYS AN IMPORTANT ROLE IN ACUTE CELLULAR ALLOGRAFT			
CC	REJECTION AND PROBABLY ALSO IN XENOGRRAFT REJECTION.			
CC	-1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.			
CC	-1- SIMILARITY: TO OTHER SELECTIN/LIBCAM.			
CC	-1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.			
CC	-1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.			
CC	-1- SIMILARITY: CONTAINS 4 SUSHI (SCR) REPEATS; PORCINE E-LECTIN LACKS			
CC	THE HUMAN SUSHI-1 AND -4 EQUIVALENTS.			
CC	-----			
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			

```
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L39076; AAA61545.1; -
DR EMBL; U08350; AAA21541.1; -
DR HSSP; P16581; 1EST.
DR PRINTS; PRO00343; SELECTIN.
DR PROSITE; PS00032; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00615; C_TYPE_LLECTIN_1; 1.
DR PROSITE; PS50041; C_TYPE_LLECTIN_2; 1.
DR PFAM; PF00008; EGF; 1.
DR PFAM; PF00059; lectin_c; 1.
DR PFAM; PF00084; sushi; 4.
KW Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;
KW Selectin; signal; Sushi; Repeat.
FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 484 E-SELECTIN.
FT TRANSMEM 23 429 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 430 451 POTENTIAL.
FT DOMAIN 452 484 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 39 139 C-TYPE LECTIN (SHORT FORM).
FT DOMAIN 140 176 EGF-LIKE.
FT REPEAT 180 421 4 X SUSHI (SCR) REPEATS.
FT REPEAT 180 236 SUSHI 1.
FT REPEAT 229 299 SUSHI 2.
FT REPEAT 302 362 SUSHI 3.
FT REPEAT 365 421 SUSHI 4.
FT DISULFID 41 139 BY SIMILARITY.
FT DISULFID 112 131 BY SIMILARITY.
FT DISULFID 144 155 BY SIMILARITY.
FT DISULFID 149 164 BY SIMILARITY.
FT DISULFID 166 175 BY SIMILARITY.
FT DISULFID 181 222 BY SIMILARITY.
FT DISULFID 208 235 BY SIMILARITY.
FT DISULFID 240 285 BY SIMILARITY.
FT DISULFID 271 298 BY SIMILARITY.
FT DISULFID 303 348 BY SIMILARITY.
FT DISULFID 334 361 BY SIMILARITY.
FT DISULFID 366 407 BY SIMILARITY.
FT DISULFID 393 420 BY SIMILARITY.
FT CARBOHYD 61 61 POTENTIAL.
FT CARBOHYD 65 65 POTENTIAL.
FT CARBOHYD 79 79 POTENTIAL.
FT CARBOHYD 160 160 POTENTIAL.
FT CARBOHYD 201 201 POTENTIAL.
FT CARBOHYD 254 254 POTENTIAL.
FT CARBOHYD 376 376 POTENTIAL.
FT CARBOHYD 400 400 POTENTIAL.
FT CONFLICT 253 253 C -> Y (IN REF. 2).
FT CONFLICT 313 313 L -> F (IN REF. 2).
FT CONFLICT 321 321 T -> N (IN REF. 2).
FT CONFLICT 327 327 K -> N (IN REF. 2).
FT CONFLICT 363 363 V -> A (IN REF. 2).
FT CONFLICT 384 384 V -> M (IN REF. 2).
FT CONFLICT 461 484 KEVPSSTSCLOPNKSTQMPDLI ->
NLEPRAPNAENPMPTCLLT (IN REF. 2).
AF AF47E25C1FD013 CRC64.
SQ SEQUENCE 484 AA; 52567 MM; 19667 MD; 19667 MS; 19667 MT; 19667 MU; 19667 MV; 19667 MW; 19667 MX; 19667 MY; 19667 MZ; 19667 NA; 19667 NB; 19667 NC; 19667 ND; 19667 NE; 19667 NF; 19667 NG; 19667 NH; 19667 NI; 19667 NJ; 19667 NK; 19667 NL; 19667 NM; 19667 NN; 19667 NO; 19667 NP; 19667 NQ; 19667 NR; 19667 NS; 19667 NT; 19667 NU; 19667 NV; 19667 NW; 19667 NX; 19667 NY; 19667 NZ; 19667 OA; 19667 OB; 19667 OC; 19667 OD; 19667 OE; 19667 OF; 19667 OG; 19667 OH; 19667 OI; 19667 OJ; 19667 OK; 19667 OL; 19667 OM; 19667 ON; 19667 OO; 19667 OP; 19667 OQ; 19667 OR; 19667 OS; 19667 OT; 19667 OU; 19667 OV; 19667 OW; 19667 OX; 19667 OY; 19667 OZ; 19667 PA; 19667 PB; 19667 PC; 19667 PD; 19667 PE; 19667 PF; 19667 PG; 19667 PH; 19667 PI; 19667 PJ; 19667 PK; 19667 PL; 19667 PM; 19667 PN; 19667 PO; 19667 PP; 19667 PQ; 19667 PR; 19667 PS; 19667 PT; 19667 PU; 19667 PV; 19667 PW; 19667 PX; 19667 PY; 19667 PZ; 19667 QA; 19667 QB; 19667 QC; 19667 QD; 19667 QE; 19667 QF; 19667 QG; 19667 QH; 19667 QI; 19667 QJ; 19667 QK; 19667 QL; 19667 QM; 19667 QN; 19667 QO; 19667 QQ; 19667 QR; 19667 QS; 19667 QT; 19667 QU; 19667 QV; 19667 QW; 19667 QX; 19667 QY; 19667 QZ; 19667 RA; 19667 RB; 19667 RC; 19667 RD; 19667 RE; 19667 RF; 19667 RG; 19667 RH; 19667 RI; 19667 RJ; 19667 RK; 19667 RL; 19667 RM; 19667 RN; 19667 RO; 19667 RP; 19667 RQ; 19667 RR; 19667 RS; 19667 RT; 19667 RU; 19667 RV; 19667 RW; 19667 RX; 19667 RY; 19667 RZ; 19667 SA; 19667 SB; 19667 SC; 19667 SD; 19667 SE; 19667 SF; 19667 SG; 19667 SH; 19667 SI; 19667 SJ; 19667 SK; 19667 SL; 19667 SM; 19667 SN; 19667 SO; 19667 SP; 19667 SQ; 19667 SR; 19667 SS; 19667 ST; 19667 SU; 19667 SV; 19667 SW; 19667 SX; 19667 SY; 19667 SZ; 19667 TA; 19667 TB; 19667 TC; 19667 TD; 19667 TE; 19667 TF; 19667 TG; 19667 TH; 19667 TI; 19667 TJ; 19667 TK; 19667 TL; 19667 TM; 19667 TN; 19667 TO; 19667 TP; 19667 TQ; 19667 TR; 19667 TS; 19667 TT; 19667 TU; 19667 TV; 19667 TW; 19667 TX; 19667 TY; 19667 TZ; 19667 UA; 19667 UB; 19667 UC; 19667 UD; 19667 UE; 19667 UF; 19667 UG; 19667 UH; 19667 UI; 19667 UJ; 19667 UK; 19667 UL; 19667 UM; 19667 UN; 19667 UO; 19667 UP; 19667 UQ; 19667 UR; 19667 US; 19667 UT; 19667 UU; 19667 UV; 19667 UW; 19667 UX; 19667 UY; 19667 UZ; 19667 VA; 19667 VB; 19667 VC; 19667 VD; 19667 VE; 19667 VF; 19667 VG; 19667 VH; 19667 VI; 19667 VJ; 19667 VK; 19667 VL; 19667 VM; 19667 VN; 19667 VO; 19667 VP; 19667 VQ; 19667 VR; 19667 VS; 19667 VT; 19667 VU; 19667 VW; 19667 VX; 19667 VY; 19667 VZ; 19667 WA; 19667 WB; 19667 WC; 19667 WD; 19667 WE; 19667 WF; 19667 WG; 19667 WH; 19667 WI; 19667 WJ; 19667 WK; 19667 WL; 19667 WM; 19667 WN; 19667 WO; 19667 WP; 19667 WQ; 19667 WR; 19667 WS; 19667 WT; 19667 WU; 19667 WV; 19667 WW; 19667 WX; 19667 WY; 19667 WZ; 19667 XX; 19667 XY; 19667 XZ; 19667 YA; 19667 YB; 19667 YC; 19667 YD; 19667 YE; 19667 YF; 19667 YG; 19667 YH; 19667 YI; 19667 YJ; 19667 YK; 19667 YL; 19667 YM; 19667 YN; 19667 YO; 19667 YP; 19667 YQ; 19667 YR; 19667 YS; 19667 YT; 19667 YU; 19667 YV; 19667 YW; 19667 YX; 19667 YY; 19667 YZ; 19667 ZA; 19667 ZB; 19667 ZC; 19667 ZD; 19667 ZE; 19667 ZF; 19667 ZG; 19667 ZH; 19667 ZI; 19667 ZJ; 19667 ZK; 19667 ZL; 19667 ZM; 19667 ZN; 19667 ZO; 19667 ZP; 19667 ZQ; 19667 ZR; 19667 ZS; 19667 ZT; 19667 ZU; 19667 ZV; 19667 ZW; 19667 ZX; 19667 ZY; 19667 ZZ.
Query Match 19.6%; Score 308; DB 1; Length 484;
Best Local Similarity 28.9%; Pred. No. 8,32e-54;
Matches 52; Conservative 38; Mismatches 82; Indels 8; Gaps 6;
```


Query Match	Best Local Similarity	Score	DB 1:	Length	768:
Matches 49;	Conservative	33;	Mismatches 75;	Indels 10;	Gaps 10;
290	FECOPGYRMRSDILHCTDSGOW-S-EPLPLTCEAIAC-EPLESPHLSMOCFPGTGAFCG	346			
32	YKCRPGYRSLGNVIMVCRK-GEWALNLPKRCQKRPCHGPDYF-GTFT-LTGVNVEY	88			
347	NSSCFETCTEGEVLVMDAIIH-CADLQWTAFAVCEALOCOEFPVPSKAQ-VSCS-DPF	403			
89	GKVAAYTCNEGYQLLGEINVREC-DTIDGMINDIPICVAVCLPTAPAEENKIIYSNAEPD	147			

QY	148	REYHFGQAFRCVCSNGYKIEGDEEMHCSDDFGWSKXPKCVETISCKS	194
RESULT	8		
ID	AP0H_MOUSE	STANDARD:	PRT; 345 AA.
AC	001339:		
DT	01-APR-1993 (Rel. 25, Created)		
DT	01-APR-1993 (Rel. 25, Last sequence update)		
DT	15-JUL-1999 (Rel. 38, Last annotation update)		
DE	BETA-2-GLYCOPROTEIN I PRECURSOR (APOLIPOPROTEIN H) (APO-H) (ACTIVATED		
DE	PROTEIN C-BINDING PROTEIN) (APC INHIBITOR) (B2GP1).		
GN	AP0H OR B2GP1.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;		
OC	Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	MEDLINE: 92372000.		
RT	Noraka M., Matsuda Y., Shirotshi T., Moriwak K., Natsume-Sakai S.;		
RT	"Molecular cloning of mouse beta 2-glycoprotein I and mapping of the		
RT	gene to chromosome 11.";		
RL	Genomics 13:1082-1087(1992).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=CBA/J; TISSUE=LIVER;		
RX	MEDLINE: 94242017.		
RA	Sellar G.C., Steel D.M., Zafirooulos A., Seery L.T.,		
RA	Whitehead A.S.;		
RT	"Characterization, expression and evolution of mouse beta 2-		
RT	glycoprotein I (apolipoprotein H) ";		
RL	Biochem. Biophys. Res. Commun. 200:1521-1528(1994).		
RN	[3]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=BALB/C; TISSUE=LIVER;		
RA	Kristensen T.;		
RT	"Structure of the human beta-2-glycoprotein I gene.";		
RT	Submitted (FEB-1997) to the EMBL/Genbank/DBP databases.		
RL	-1- FUNCTION: RINDS TO VARIOUS KINDS OF PROTEIN		

SUCH AS HEPARIN, PHOSPHOLIPIDS, AND DEXTRAN SULFATE, MAY PREVENT ACTIVATION OF THE INTRINSIC BLOOD COAGULATION CASCADE BY BINDING TO PHOSPHOLIPIDS ON THE SURFACE OF DAMAGED CELLS.

-1- TISSUE SPECIFICITY: PLASMA.

-1- SIMILARITY: CONTAINS 4 SUSHI (SCR) REPEATS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb.ch/announce/> or send an email to license@isb.sib.ch).

CC

DR EMBL: D10056; BAA00945.1; -

DR EMBL: S70439; AAB30789.1; -

DR EMBL: Y11356; CAA72190.1; -

DR PIR: A43286; NEMS.

DR HSP: P10998; 1VVC.

DR MGD: MGI:88058; APOH.

DR PFAM: PF00084; sushi; 4.

KW Heparin-binding; Glycoprotein; Plasma; Repeat; Sushi; Signal.

FT SIGNAL 1 19

FT CHAIN 20 345 BETA-2-GLYCOPROTEIN I.

FT DOMAIN 22 261 4 X SUSHI (SCR) REPEATS.

FT REPEAT 22 80 SUSHI 1.

FT REPEAT 83 138 SUSHI 2.

FT REPEAT 141 201 SUSHI 3.

FT REPEAT 204 261 SUSHI 4.

FT DOMAIN 263 345 MODIFIED-SUSHI.

FT DISULFID 23 66 BY SIMILARITY.

FT DISULFID 51 79 BY SIMILARITY.

FT DISULFID 84 124 BY SIMILARITY.

FT DISULFID 110 137 BY SIMILARITY.

FT DISULFID 142 188 BY SIMILARITY.

FT DISULFID 174 200 BY SIMILARITY.

FT DISULFID 205 240 BY SIMILARITY.

FT DISULFID 234 260 BY SIMILARITY.

FT DISULFID 264 315 BY SIMILARITY.

FT DISULFID 300 325 BY SIMILARITY.

FT DISULFID 307 345 BY SIMILARITY.

FT CARBOHYD 162 162 POTENTIAL.

FT CARBOHYD 183 183 POTENTIAL.

FT CARBOHYD 193 193 POTENTIAL.

FT CONFLICT 252 252 G -> R (IN REF. 2).

SO SEQUENCE 345 AA; 38619 MM; C83F8A6BBD51C940 CRC64;

Query Match 18.9%; Score 297; DB 1; Length 345;

Best Local Similarity 29.0%; Pred. No. 7,31e-51;

Matches 51; Conservative 37; Mismatches 78; Indels 10; Gaps 8;

DB 40 SYDPEQIVYCKRGVGRGMRRFTCLTGMW-PINTLR-CVPRVCPFAGILENGIVRY 97

QY 23 TYPESTQIVYCRPEYRSLGVIM-VCR-KGEWALNLRKCRKPGCHPBDTFPGFTL 80

DB 98 TS---FEYKPNISFACNPGF-FLNCTSSSKCTEGKWSPDIPACARITCPPPVKFAALL 153

QY 81 TGGNFEYGVAVYTCNGYQLGEINVRCDTDLG-WTNDIPICGVYACLPVTAPEMNGKI 139

DB 134 KDPRSAAGNSLYQDTVYFKCLPFAHMGNDYVACTEGGNTRL-PELEVKCFPP 208

QY 140 VSSAMEPREYHFGAVRFVCSNGYKIEGDEMHCSDDGFWSKPKCVETSCASP 195

RESULT 9

ID LEM2_HUMAN STANDARD; PRT; 610 AA.

AC P16581; P16111;

DI 01-APR-1990 (Rel. 14, Created)

DI 01-AUG-1990 (Rel. 15, Last sequence update)

DI 15-JUL-1999 (Rel. 38, Last annotation update)

DE E-SELECTIN PRECURSOR (ENDOTHELIAL LEUKOCYTE ADHESION MOLECULE 1)

DE (ELAM-1) (LEUKOCYTE-ENDOTHELIAL CELL ADHESION MOLECULE 2) (LECAM2) (CD62E).

GN SELE OR ELAM1.

OS Homo sapiens (Human).

CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

CC Eutheria; Primates; Catarrhini; Homidae; Homo.

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE: 90175359.

RA Hession C., Osborn L., Goff D., Chi-Rosso G., Vassallo C., Paisek M., Pittack C., Tizack R., Goetz S., McCarthy K., Hopple S., Lobb R.;

RT "Endothelial leukocyte adhesion molecule 1: direct expression cloning and functional interactions.";

RL Proc. Natl. Acad. Sci. U.S.A. 87:1673-1677(1990).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE: 89162047.

RA Bevilacqua M.P., Stengelin S., Gimbrone M.A. Jr., Seed B.;

RT "Endothelial leukocyte adhesion molecule 1: an inducible receptor for neutrophils related to complement regulatory proteins and lectins.";

RL Science 243:1160-1165(1989).

RN [3]

RP SEQUENCE FROM N.A.

RX MEDLINE: 91115870.

RA Collins T., Williams A., Johnston G.I., Kim J., Eddy R., Shows T., Gimbrone M.A. Jr., Bevilacqua M.P.;

RT "Structure and chromosomal location of the gene for endothelial-leukocyte adhesion molecule 1.";

RL J. Biol. Chem. 266:2466-2473(1991).

RN [4]

RP LIGAND.

RX MEDLINE: 91068005.

RA Phillips M.L., Nudelmann E., Gaeta F.C., Perez M., Singhal A.K., Hakomori S., Paulson J.C.;

RT "ELAM-1 mediates cell adhesion by recognition of a carbohydrate ligand, sialyl-Lex.";

RL Science 250:1130-1132(1990).

RN [5]

RP 3D-STRUCTURE MODELING OF LECTIN DOMAIN.

RX MEDLINE: 93202275.

RA Mills A.;

RT "Modelling the carbohydrate recognition domain of human E-selectin.";

RL FEBS Lett. 319:5-11(1993).

RN [6]

RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 22-178.

RX MEDLINE: 94150646.

RA Graves B.J., Crowther R.L., Chandran C., Rumberger J.M., Li S., Huang K.-S., Presky D.H., Familletti P.C., Wolitzky B.A., Burns D.K.;

RT "Insight into E-selectin/ligand interaction from the crystal structure and mutagenesis of the lec/EGF domains.";

RL Nature 367:532-538(1994).

RN [7]

RP VARIANT ARG-149.

RX MEDLINE: 95179107.

RA Wenzel K., Felix S., Kleber F.X., Brachold R., Menke T., Schatke S., Schulte K.L., Glaser C., Rohde K., Baumann G., Speer A.;

RT "E-selectin polymorphism and atherosclerosis: an association study.";

RL Hum. Mol. Genet. 3:1935-1937(1994).

RN [8]

RP VARIANT ARG-149.

RX MEDLINE: 99134508.

RA Ye S.Q., Usher D., Virgil D., Zhang L.Q., Yochim S.E., Gupta R.;

RT "A beta polymorphism detects the mutation of serine-128 to arginine in CD 62 gene - a risk factor for coronary artery disease.";

RL J. Biomed. Sci. 6:18-21(1999).

CC -1- FUNCTION: EXPRESSED ON CYTOKINE INDUCED ENDOTHELIAL CELLS AND MEDIATES THEIR BINDING TO LEUKOCYTES. THE LIGAND RECOGNIZED BY ELAM-1 IS SIALYL-LEWIS X (ALPHA(1->3)FUCOSYLATED DERIVATIVES OF POLYACROSAMINE THAT ARE FOUND AT THE NONREDUCING TERMINI OF GLYCOPOLYIDS).

CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.

CC -1- POLYMORPHISM: A POLYMORPHISM IN POSITION 149 IS ASSOCIATED WITH A HIGHER RISK OF CORONARY ARTERY DISEASE (CAD). A SIGNIFICANTLY HIGHER MUTATION FREQUENCY (ARG-149) IS OBSERVED IN PATIENTS WITH

OY 135 ENKIVSSAMEPDREYHFGAVFVNCNGYKIEGDEMHCS-D-D--GFWSEKPKVEI 190
 Db 238 TCRPDVSHGEMVS 251
 : : : : :
 OY 191 SCKSPDVINGSPIS 204

RESULT 13
 ID CCPH.HSVSA STANDARD; PRT; 360 AA.
 AC 001016;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE COMPLEMENT CONTROL PROTEIN HOMOLOG PRECURSOR (CCPH).
 GN 4 OR CCPH.
 OS Herpesvirus saimiri (strain 11).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Gammaherpesvirinae.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 92333688.
 RA Albrecht J.-C., Nicholas J., Biller D., Cameron K.R., Biesinger B.,
 RA Newman C., Wiltman S., Craxton M.A., Coleman H., Fleckenstein B.,
 RA Honess R.W.;
 RT "Primary structure of the herpesvirus saimiri genome."
 RL J. Virol. 66:5047-5058(1992).
 RN [2]
 RP SIMILARITY TO CCP.
 RX MEDLINE: 92260674.
 RA Albrecht J.-C., Fleckenstein B.;
 RT "New member of the multigene family of complement control proteins in
 RT herpesvirus saimiri."
 RL J. Virol. 66:3937-3940(1992).
 CC -1- ALTERNATIVE PRODUCTS: A MEMBRANE-BOUND FORM AND A SECRETED FORM
 CC ARE PRODUCED BY ALTERNATIVE SPLICING OF THE SAME GENE.
 CC -1- SIMILARITY: BELONGS TO THE SUPERFAMILY OF THE REGULATORS OF
 CC COMPLEMENT ACTIVATION (RCA).
 CC -1- SIMILARITY: CONTAINS 3 SUSHI (SCR) REPEATS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: X64346; CAA5626.1; -
 DR EMBL: X64346; CAA5627.1; -
 DR EMBL: X60283; CAA42823.1; -
 DR EMBL: X60283; CAA42822.1; -
 DR PIR: BA2534; WMBE2E.
 DR PIR: BA2534; WMBE2E.
 DR PIR: S24567; S24567.
 DR HSSP: P10998; IYVC.
 DR PIR: PF00084; SUSHI; 4.
 DR KW Signal; Repeat; SUSHI; Transmembrane; Alternative splicing;
 DR KW Glycoprotein.
 FT SIGNAL 1 20 POTENTIAL.
 FT CHAIN 1 360 COMPLEMENT CONTROL PROTEIN HOMOLOG.
 FT DOMAIN 83 265 SUSHI 1.
 FT REPEAT 83 143 SUSHI 2.
 FT REPEAT 146 206 SUSHI 3.
 FT REPEAT 209 265 SUSHI 3.
 FT TRANSMEM 328 350 POTENTIAL.
 FT DISULFID 84 125 BY SIMILARITY.
 FT DISULFID 111 142 BY SIMILARITY.
 FT DISULFID 147 191 BY SIMILARITY.
 FT DISULFID 175 205 BY SIMILARITY.
 FT DISULFID 210 252 BY SIMILARITY.
 FT DISULFID 238 264 BY SIMILARITY.
 FT CARBOHYD 36 36 POTENTIAL.
 FT CARBOHYD 39 39 POTENTIAL.

FT CARBOHYD 46 46 POTENTIAL.
 FT CARBOHYD 72 72 POTENTIAL.
 FT CARBOHYD 155 155 POTENTIAL.
 FT CARBOHYD 294 294 POTENTIAL.
 FT VARSPLIC 289 302 R10NGNCTSMPTQ -> AECACGGSNYPIS (IN
 FT VARSPLIC 303 360 MISSING (IN SHORT ISOFORM).
 FT VARSPLIC 360 40006 MW; 6278A6C2ECD49669 CRG64;
 SQ SEQUENCE 360 AA; 40006 MW; 6278A6C2ECD49669 CRG64;
 Query Match 18.3%; Score 288; DB 1; Length 360;
 Best Local Similarity 34.68; Pred. 1.82e-48;
 Matches 64; Conservative 33; Mismatches 70; Indels 18; Gaps 13;

Db 44 YPNGTLLHVCRCRGYARPPQVITVCVGNM-TV-P-KKCKKCKSPPODLNGRYVTG- 99
 : : : : :
 OY 24 YPEGTQALVYKCRGYSRLGVMVNCVRCGEVVALNPLKCKRCGCHGDTPTFTLTGG 83
 : : : : :
 Db 100 NLV-YGSVITYTNGSYSLGTT-SACLKRGKRVDMTPRPICDKCKP--PQIAN 155
 : : : : :
 OY 84 NVEYGVKAVYTCNEGQYLGEINREK--DTDG--WTNDIPICEVVKCLPTAVENGR 138
 : : : : :
 Db 156 GTHDNVYK-DE-YTYLDVYVYSCNDEKRLTLTGSSKLCSTGSGWVNGETKCEFIKCLP 213
 : : : : :
 OY 139 IVSSAMEPDREYHFGAVFVNCNGYKIE--GDEMHCSDDGFW-SKERPKVEISKSP 195
 : : : : :
 Db 214 QVANA 218
 : : : : :
 OY 196 DVING 200

RESULT 14
 ID LEM2.MOUSE STANDARD; PRT; 612 AA.
 AC 000690;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE E-SELECTIN PRECURSOR (ENDOTHELIAL LEUKOCYTE ADHESION MOLECULE 1)
 DE (ELAM-1) (LEUKOCYTE-ENDOTHELIAL CELL ADHESION MOLECULE 2) (LECAM2)
 DE (CD62E).
 GN SELE OR ELAM-1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia;
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 92283265.
 RA Becker-Andre M., van Huijsduijn R.H., Losberger C., Whelan J.,
 RA Delamarier J.F.;
 RT "Murine endothelial leukocyte-adhesion molecule 1 is a close
 RT structural and functional homologue of the human protein.";
 RL Eur. J. Biochem. 206:401-411(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 92340571.
 RA Weller A., Isenmann S., Vestweber D.;
 RT "Cloning of the mouse endothelial selectins. Expression of both E-
 RT and P-selectin is inducible by tumor necrosis factor alpha.";
 RL J. Biol. Chem. 267:15176-15183(1992).
 CC -1- FUNCTION: EXPRESSED ON CYTOKINE INDUCED ENDOTHELIAL CELLS AND
 CC MEDIATES THEIR BINDING TO LEUKOCYTES. THE LIGAND RECOGNIZED BY
 CC POLYLACTOSAMINE THAT ARE FOUND AT THE NONREDUCING TERMINI OF
 CC GLYCOPOLYDPS).
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -1- SIMILARITY: TO OTHER SELECTINS/LECAMs.
 CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
 CC -1- SIMILARITY: CONTAINS 6 SUSHI (SCR) REPEATS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

[illegible]

QY	118	NDIICEVWKCPLPTAFENGRKIVSNAEPDREHFQGAIVFYVNCGKXIGDEMHCSDD	177
Db	417	GEMDSKRPTCSAVKC	431
QY	178	GFWSEKPKVCVEISC	192
RESULT	15		
ID	LEM2_RAT	STANDARD;	PRT; 549 AA.
AC	P98105;		
DT	01-FEB-1996 (Rel. 33, Created)		
DT	01-FEB-1996 (Rel. 33, Last sequence update)		
DT	01-FEB-1996 (Rel. 33, Last annotation update)		
DE	E-SELECTIN PRECURSOR (ENDOTHELIAL LEUKOCYTE ADHESION MOLECULE 1)		
DE	(ELAM-1) (LEUKOCYTE-ENDOTHELIAL CELL ADHESION MOLECULE 2) (LECAM2)		
DE	(CD62E).		
GN	SELE OR ELAM-1.		
OS	Rattus norvegicus (Rat).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;		
OC	Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=LUNG;		
RA	Rosenbloom C.L., Auchampach J.A., Anderson D.C., Manning A.M.;		
RL	Submitted (NOV-1993) to the EMBL/Genbank/DBJ databases.		
CC	-1- FUNCTION: EXPRESSED ON CYTOKINE INDUCED ENDOTHELIAL CELLS AND		
CC	MEDIATES THEIR BINDING TO LEUKOCYTES. THE LIGAND RECOGNIZED BY		
CC	ELAM-1 IS SIMILY-LEWIS X (ALPHA1-3)FUCOSYLATED DERIVATIVES OF		
CC	POLYACTOSAMINE THAT ARE FOUND AT THE NONREDUCING TERMINI OF		
CC	GLYCOLIPIDS).		
CC	-1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.		
CC	-1- SIMILARITY: TO OTHER SELECTINS/LECAMS.		
CC	-1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.		
CC	-1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.		
CC	-1- SIMILARITY: CONTAINS 5 SUSHI (SCR) REPEATS.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/		
CC	or send an email to license@isb-sib.ch).		
CC	-----		
CC	EMBL; L25527; AAA41113.1; .		
DR	HSSP; P16581; IKJA.		
DR	PROSITE; PS00022; EGF_1; 1.		
DR	PROSITE; PS01186; EGF_2; 1.		
DR	PROSITE; PS00615; C-TYPE LECTIN_1; 1.		
DR	PROSITE; PS50041; C-TYPE LECTIN_2; 1.		
DR	PFAM; PF00008; EGF; 1.		
DR	PFAM; PF00059; lectin_c; 1.		
DR	PFAM; PF00084; sushi; 5.		
KW	Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;		
KW	Selectin; Signal; Sushi; Repeat.		
FT	SIGNAL	1	21
FT	CHAIN	22	549
FT	DOMAIN	22	494
FT	TRANSMEM	495	516
FT	DOMAIN	517	549
FT	DOMAIN	38	138
FT	DOMAIN	139	175
FT	DOMAIN	179	486
FT	DOMAIN	179	239
FT	REPEAT	242	301
FT	REPEAT	304	364
FT	REPEAT	367	427
FT	REPEAT	430	486
FT	DISULFID	40	138
FT	DISULFID	111	130
FT	DISULFID	143	154
FT	DISULFID	148	163
FT	DISULFID	148	163

```

FT DISULFID 165 174 BY SIMILARITY.
FT DISULFID 180 225 BY SIMILARITY.
FT DISULFID 210 238 BY SIMILARITY.
FT DISULFID 243 287 BY SIMILARITY.
FT DISULFID 273 300 BY SIMILARITY.
FT DISULFID 305 350 BY SIMILARITY.
FT DISULFID 336 363 BY SIMILARITY.
FT DISULFID 368 413 BY SIMILARITY.
FT DISULFID 399 426 BY SIMILARITY.
FT DISULFID 431 472 BY SIMILARITY.
FT DISULFID 458 485 BY SIMILARITY.
FT CARBOHYD 25 25 POTENTIAL.
FT CARBOHYD 60 60 POTENTIAL.
FT CARBOHYD 145 145 POTENTIAL.
FT CARBOHYD 192 192 POTENTIAL.
FT CARBOHYD 203 203 POTENTIAL.
FT CARBOHYD 266 266 POTENTIAL.
FT CARBOHYD 313 313 POTENTIAL.
FT CARBOHYD 320 330 POTENTIAL.
FT CARBOHYD 333 333 POTENTIAL.
FT CARBOHYD 441 441 POTENTIAL.
FT CARBOHYD 465 465 POTENTIAL.
SQ SEQUENCE 549 AA: 60079 MW: 85CECD87B0144C8 CRC64;

```

Query Match 18.2% Score 286; DB 1; Length 549;
 Best Local Similarity 25.1%; Pred. No. 6.17e-48;
 Matches 50; Conservative 45; Mismatches 97; Indels 7; Gaps 7;

```

DB 242 ECKALTPAHG-VKCCSSNPGSYPMWTTCTFDCEGRRVGAONLOCTSSGVWDNEKPSG 300
QY 2 DNELEPPRRNTIELTGSWSDQTYPEGTQAIKCRPGYRSIGNYIMVC-RKGEWVALNPLR 60
DB 301 KAVTCDAIRPONGSVSCSNSTAGAL-AFKSSCNFTCEHSTFLQSPAQV-ECSAQCQWTP 358
QY 61 KQOK-RPGCHPDGDTFRGFTTLTGANVFEGYKAVYTCNEGYYDLGEINRYECDTDG-WTN 118
DB 359 QIPVCKASOCALSAPOQGHMKCLPSASAP-FQSGSSCKFSCDEGPELKGSRRLQCGPRG 417
QY 119 DIPICEVVKCLPFTAPENCKIVSAMPEPDREYHFGQAVRFTVCSGKIKIDGDEMHCSDDG 178
DB 418 EWDSEKPTCAGVQCSSDL 436
QY 179 FWSKEKPKCVETSCSPDV 197

```

Search completed: Thu Jun 8 21:38:24 2000
 Job time : 12 secs.

```

Mpsrch_pp    protein - protein database search, using Smith-Waterman algorithm
Run on:      Thu Jun  8 21:37:08 2000;      MasPar time 8.38 Seconds
Tabular output not generated.              585,300 Million cell updates/sec

```

Scoring table: PAM 150
Gap 11

Searched: 188963 seqs, 23686106 residues

Post-processing: Minimum Match 0%
Listing First 45 summaries

Database: a-~~genes~~seq36

Statistics: Mean 31.247; Variance 114.255; scale 0.273

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	973	61.9	240	1	W39154	Human partial Complement	6.80e-97
2	842	53.5	216	1	W39155	Clone pRR99FHA10 CFH r	8.32e-82
3	308	19.6	484	1	W93692	Porcine E-selectin.	1.19e-21
4	308	19.6	830	1	R65216	P-selectin.	1.19e-21
5	298	18.9	610	1	R08116	Endothelial cell-leuco	1.43e-20
6	298	18.9	610	1	R05494	Endothelial leukocyte	1.43e-20
7	298	18.9	610	1	R05538	Endothelial-leukocyte	1.43e-20
8	298	18.9	610	1	W18839	E-selectin.	1.43e-20
9	298	18.9	610	1	W45733	Amino acid sequence of	1.43e-20
10	293	18.6	259	1	W06881	Decay accelerating fac	4.96e-20
11	293	18.6	381	1	W26317	Human decay acceleratrl	4.96e-20
12	293	18.6	381	1	R66683	Decay accelerating fac	4.96e-20
13	293	18.6	381	1	P70048	Human decay acceleratrl	4.96e-20
14	293	18.6	381	1	P94773	Decay accelerating fac	4.96e-20
15	293	18.6	381	1	W73505	Decay accelerating fac	4.96e-20
16	293	18.6	440	1	P94774	Membrane bound decay a	4.96e-20
17	293	18.6	440	1	W27483	Human glycoposphatidy	4.96e-20
18	293	18.6	440	1	P70049	Human decay acceleratrl	4.96e-20
19	293	18.6	440	1	R66684	Membrane accelerating fac	4.96e-20
20	293	18.6	577	1	W06882	Decay co-factor pro	1.04e-19
21	290	18.4	581	1	R13490	Human C4 binding prote	1.71e-19
22	288	18.3	302	1	W26370	Hepesvirus secreted g	1.71e-19
23	288	18.3	302	1	R57973	Herpesvirus salmimrl sc	1.71e-19

ALIGNMENTS

RESULT	1
ID	W39154 standard; Protein; 240 AA.

DT 27-APR-1998. (first entry)
DE Human partial Complement factor H protein fragment 1.
KW Complement factor H; tumour associated antigen; renal cancer;
KW urogenital cancer; mediastinum; modulator.
OS Homo sapiens.
OS W09738156-A1.
PN 16-OCT-1897.
PD 16-OCT-1897.
PF 09-APR-1997; 005710.
PR 06-MAR-1997; US-812481.
PR 09-APR-1996; US-015083.
PR 09-APR-1996; US-630048.
PR 06-MAR-1997; US-038614.
PA (BARO-) BARO DIAGNOSTIC SCI INC.
PI Enfield IL, Hass GM, Kinders RJ;
PI WPI: 97-512742/47.
DR N-BSDs: V02790.
PT Treating or screening for cancer, e.g. renal or urogenital cancer -
PT by modulating or detecting tumour associated human complement factor
PT H related antigen, or nucleic acid encoding it
PS Example 6b; Fig 6b; 104pp; English. Abstracts a review of the human

PS Example 1b) A protein sequence represents a region of the human
CC This partial protein sequence represents a region of the human
CC tumour associated complement factor H (CFH). This sequence is used
CC in the identification of complement factor H related proteins and
CC antigens isolated from clone PB89F.H10 (see W93155). The detection of
CC such proteins and a CFH antigens can be used in screening or for the
CC treatment of renal or urogenital cancer, e.g. bladder, cervical or
CC prostate cancer. Agents that may modulate this antigen could be used in
CC the manufacture of a medicament for the treatment of a tumour cell.
SQ Sequence 240 AA:

Query Match	61.9%	Score 973	DB 1	Length 240
Best Local Similarity	100.0%	Pred. No. 6.80e-97		
Matches 130	Conservative	0	Mismatches 0	Indels 0
			Gaps	0

Db	1	FTLTGWNFEYGYKAVYTTNEGYYOLLGELINRBCDTGDMTNDIPICEVVKCLPTAPANG	60
Qy	78	FTLTGWNFEYGYKAVYTTNEGYYOLLGELINRBCDTGDMTNDIPICEVVKCLPTAPANG	137
Db	61	KIYSSAMEPRREYHFGQAVRFVCSNSGKRIEGDEMHKSDSGFWSKRPKCVIEICKSPDV	120
Qy	138	KIYSSAMEPRREYHFGQAVRFVCSNSGKRIEGDEMHKSDSGFWSKRPKCVIEICKSPDV	197
Db	121	INGSPISQKI 130	

QY 198 INGSPIISOKI 207

RESULT

ID W39155 standard; Protein; 216 AA.

AC W39155; 27-APR-1998 (first entry)

DE Clone PRB9FH410 CFH related protein fragment.

KM Complement factor H; tumour associated antigen; renal cancer;

OS urogenital cancer; modulator.

PD W09738136-A1.

PF 16-OCT-1997.

PR 09-APR-1997; 005710.

PR 06-MAR-1997; US-812481.

PR 09-APR-1996; US-015083.

PR 06-MAR-1997; US-038614.

PA (BARD) BARD DIAGNOSTIC SCI INC.

PI Enfield DL, Haas GM, Kinders RJ;

DR WPI; 97-512742/47.

DR N-PSDB; V02791.

PT Treating or screening for cancer, e.g. renal or urogenital cancer -

PT by modulating or detecting tumour associated human complement factor

PS Example 6B; Fig 6B; 104pp; English.

CC This partial protein is found in clone PRB9FH410 and represents a

CC complement factor H related protein with homology to a region of the

CC human tumour-associated complement factor H (CFH). The detection of this

CC protein and a CFH antigen can be used in screening for or the treatment

CC of renal or urogenital cancer, e.g. bladder, cervical or prostate cancer.

CC Agents that may modulate this antigen could be used in the manufacture of

CC a medicament for the treatment of a tumour cell.

SQ Sequence 216 AA.

Query Match Best Local Similarity 53.5%; Score 842; DB 1; Length 216;

Matches 112; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 1 TCNEGYQLGELNRYRCDGTNDIPICVYKCLPTAPENKGVSSAMEPDRYHFGQ 60

QY 95 TCNEGYQLGELNRYRCDGTNDIPICVYKCLPTAPENKGVSSAMEPDRYHFGQ 154

Db 61 AVRFVNCNGYKIEGDEHMCSDGFGWKEKPKVCVEISCKSPVINGSPISOKI 113

QY 155 AVRFVNCNGYKIEGDEHMCSDGFGWKEKPKVCVEISCKSPVINGSPISOKI 207

RESULT 3

ID W99892 standard; Protein; 484 AA.

AC W99892; 10-JUN-1999 (first entry)

DE Porcine E-selectin.

KM Porcine; E-selectin; diagnosis; transplant rejection; xenotransplant;

KM cytokine; donor organ endothelium; inflammatory cell; graft tissue;

KM acute cellular allograft rejection; cell adhesion.

OS sue scrofa.

PN US891645-A.

PD 06-APR-1999.

PR 01-JUN-1994; 252493.

PR 01-JUN-1994; US-252493.

PA (ALEX-) ALEXON PHARM INC.

PI Evans MJ, Matlis LA, Rollins S, Rother RP;

DR WPI; 99-253853/21.

DR N-PSDB; X19809.

PT Nucleic acids encoding the porcine E-selectin protein - useful in

PT diagnosing and treating rejection of xenotransplanted pig organs

PS Claim 1; Column 47-50; 30pp; English.

CC The present sequence is porcine E-selectin (E-sel). The nucleotide

CC sequence encoding E-sel may be used in the diagnosis, prevention and

CC treatment of tissue rejection following xenotransplants of pig organs

CC and also in the monitoring (by specific measurement of the amount of

CC E-sel in the blood of the patient) and modulation of immune responses to

CC the transplants. Cytokine-induced expression of E-sel by donor organ
 CC endothelium, contributes to the binding and subsequent transmigration of
 CC inflammatory cells into the graft tissue that occurs during acute
 CC cellular allograft rejection. Therefore, preventing expression of E-sel
 CC on the donor tissue may reduce the risk of rejection by modulating E-sel
 CC mediated cell adhesion. Fragments of the nucleotide sequence encoding
 CC hybridization probes for the identification and/or isolation of the
 CC porcine E-sel gene from genomic DNA. Antisense sequences may be used to
 CC inhibit the expression of E-sel. DNA constructs may be used to introduce
 CC antisense sequences to a cell, or to inactivate the E-sel gene (e.g.
 CC knock-out constructs). Cells in which the expression of E-sel has been
 CC prevented may be used as disease models or to produce transgenic animals
 CC (e.g. pigs) from which organs may be harvested for transplantation with
 CC a reduced risk of rejection.

SQ Sequence 484 AA;

Query Match Best Local Similarity 19.6%; Score 308; DB 1; Length 484;

Matches 52; Conservative 38; Mismatches 82; Indels 8; Gaps 6;

Db 196 QSLFWNTCAFECKEGFELIGPEHLIOCTSSGSMGKKP--TKAVNCDYGHQNDVSC 253

QY 22 QTEPGTQAIYKCRPGRSGNVIMVC-RKGEVVALNPLKCKRPGAGDPPTFTTL 80

Db 254 NHSSIGFAVYKSTCHFCAGGFGIQAQI-ECTAAGQWTOAPVCAVYKCPAVSOPKNG 312

QY 81 TGVNVEFGVAAV--YTCNNGYQLGELNRYRCDGTG-WTNDIPICVYKCLPTAPENG 137

Db 313 -LVFTHSPTEGFEYKSCAFSEEGFELRGSALACTSGGQWTOVPCVQVQCSLEY 371

QY 138 KIVSASMEPREHFGQAVRFVNCNGYKIEGDEHMCSDGFGWKEKPKVCVEISCKSPV 197

RESULT

ID R65216 standard; Protein; 830 AA.

AC R65216; 04-OCT-1995 (first entry)

DE P-selectin.

KM Control elements; gene expression; endothelial cells;

KM megakaryocytes; rheumatoid arthritis; ischaemic injury;

KM atherosclerosis; bacterial sepsis; tumour metastases.

OS Homo sapiens.

PN W09506118-A.

PD 02-MAR-1995.

PR 19-AUG-1994; U09395.

PR 20-AUG-1993; US-110158.

PA (OKLA) UNIV OKLAHOMA STATE.

PI McEver RP, Pan J;

DR WPI; 95-106847/14.

PT Expression control elements in the 5'-flanking region of the

PT P-selectin gene - for regulating gene expression in endothelial

PS cells and mega:karyocyte(s).

PT Disclosure; Fig 1; 88pp; English.

CC The sequence is that of the P-selectin gene product from endothelial

CC cells, determined from a composite of four overlapping cDNAs: lambda

CC GMP1-lambda GMP4. The P-selectin gene or its fragments may be used

CC to inhibit P-selectin expression, thus controlling inflammatory and

CC haemostatic processes in e.g. rheumatoid arthritis, ischaemic injury,

CC bacterial sepsis, tumour metastases, atherosclerosis, etc.

SQ Sequence 830 AA;

Query Match Best Local Similarity 19.6%; Score 308; DB 1; Length 830;

Matches 55; Conservative 44; Mismatches 94; Indels 12; Gaps 12;

Db 199 ECGEELPOHV-LIMNCSPILNPFNSGCSCHTGDGVNPSLCEIASGIWTN-RP-P 255

QY 2 DCELEPPRRNIEILGWSDDGTPEGTQAIYKCRPGRSGNVIMVC-RKGEVVALNPLR 60

Db 256 OCLAQC-PPKIDBERGNMILHSAKAFHOSSCSFCEGFAVVG-EEVOCASGVWT 313

QY 61 KQKRPCCGHPDTP-FGFT-LTGVNVEFGVAAVYTCNNGYQLGELNRYRCDGTG-WT 117

Db 210 CDRGYLPSSMETMOCSSGEMSA--PIPACNVECDANTNPANGFVECFQNGSPFWMTT 267
 QY 34 CRFGYSLNINIVWCRK-GEWVALNPLRKCQKRCRPGHDPDTFGFTT-LTGNNVFEYGVK 91
 Db 268 CTFDEGEFELMGAQSL-QCTSSGNDNNEKPTCKAVTCRAVROPONGS-VRCSSHAPAGEF 325
 QY 92 AYTTCNEGQYLLGEINRECDTDG-WTNDIPICEVVKCLPVTAPENCKIYSSAMEPDREY 150
 Db 326 TFKSSCNFTCEBGFMLQCPAOCVECTTQOGWTOQIPVCEAFQCTA 369
 QY 151 HFGQAVRFVNCNSGYKIEGDEMHCSDDGFWSKERKPCVEISCKS 194

RESULT 8
 ID W18839 standard; Protein: 610 AA.
 AC W18839;
 DT 05-JAN-1998 (first entry)
 DE E-selectin.
 KW E-selectin; endothelial leukocyte adhesion molecule-1; ELAM-1;
 KW murine monoclonal antibody; H4/18; H18/7; immunconjugate;
 KW activated endothelium; treat; inflammation; microbial infection;
 KW post-reperfusion injury; malignant tumour; vasculitis;
 KW vascular smooth-muscle cell proliferative disorder; acute;
 KW chronic allograft rejection.
 OS Homo sapiens.
 PN US5632991-A.
 PD 27-MAY-1997.
 PF 14-NOV-1988; 270860.
 PR 29-DEC-1994; US-365470.
 PR 14-NOV-1988; US-270860.
 PR 13-MAR-1992; US-850802.
 PR 05-AUG-1993; US-102510.
 PA (BGHM) BRIGHAM & WOMENS HOSPITAL.
 PI Gimbrone MA;
 DR WPI: 97-297321/27.
 DR N-PSDB: T70493.
 PT Immunoconjugates based on monoclonal antibody specific for
 PT E-selectin - for targeting therapeutic agent to activated
 PT endothelium
 PS Example 2: Column 35-38; 36pp: English.
 CC This sequence is E-selectin (formerly known as endothelial
 CC leukocyte adhesion molecule-1; ELAM-1). E-selectin is identified by
 CC specific binding of murine monoclonal antibodies H4/18 and H18/7
 CC developed against stimulated endothelial cells. New immunoconjugates
 CC of the invention comprise monoclonal antibody H18/7, or a fragment
 CC which binds to E-selectin conjugated to a therapeutic agent. The
 CC immunoconjugate is used to selectively target the therapeutic agent
 CC to activated endothelium expressing E-selectin, especially for treating
 CC inflammation, microbial infections, post-reperfusion injury, malignant
 CC tumours, vasculitis, vascular smooth-muscle cell proliferative disorders
 CC or acute or chronic allograft rejection.
 SQ Sequence 610 AA;

Query Match 18.9%; Score 298; DB 1; Length 610;
 Best Local Similarity 27.4%; Pred. No. 1,43e-20;
 Matches 45; Conservative 33; Mismatches 79; Indels 7; Gaps 6;
 Db 210 CDRGYLPSSMETMOCSSGEMSA--PIPACNVECDANTNPANGFVECFQNGSPFWMTT 267
 QY 34 CRFGYSLNINIVWCRK-GEWVALNPLRKCQKRCRPGHDPDTFGFTT-LTGNNVFEYGVK 91
 Db 268 CTFDEGEFELMGAQSL-QCTSSGNDNNEKPTCKAVTCRAVROPONGS-VRCSSHAPAGEF 325
 QY 92 AYTTCNEGQYLLGEINRECDTDG-WTNDIPICEVVKCLPVTAPENCKIYSSAMEPDREY 150
 Db 326 TFKSSCNFTCEBGFMLQCPAOCVECTTQOGWTOQIPVCEAFQCTA 369
 QY 151 HFGQAVRFVNCNSGYKIEGDEMHCSDDGFWSKERKPCVEISCKS 194

DT 02-JUN-1998 (first entry)
 DE Amino acid sequence of endothelial leukocyte adhesion molecule-1.
 KW Endothelial leukocyte adhesion molecule-1; ELAM-1; inducible;
 KW cytokine; overexpression; cellular adhesion molecule;
 KW intracellular adhesion molecule-1; ICAM-1; antisense molecule;
 KW lipid mixture; Alzheimer's disease; multiple sclerosis;
 KW viral hepatitis; cholangitis; cardiac allograft rejection.
 OS Homo sapiens.
 PN W0974671-AL.
 PD 11-DEC-1997.
 PF 22-MAY-1997; CA0347.
 PR 30-MAY-1996; US-657753.
 PA (UYBR) UNIV BRITISH COLUMBIA.
 PI Hope MJ, Klimuk SK, Scherrer P, Sample SC;
 DR WPI: 98-042180/04.
 DR N-PSDB: V16208.
 PT Composition for treatment of conditions associated with
 PT overexpression of ICAM-1 - used to treat e.g. Alzheimer's disease,
 PT glomerulonephritis, rheumatoid arthritis etc.
 PS Disclosure: Pages 44-46; 81pp: English.
 CC The present sequence represents human endothelial leukocyte adhesion
 CC molecule-1 (ELAM-1). ELAM-1 is a membrane glycoprotein which is induced
 CC by a number of cytokines. Pathological conditions associated with the
 CC overexpression of cellular adhesion molecules, such as ELAM-1 or
 CC intracellular adhesion molecule-1 (ICAM-1), can be treated with the
 CC novel pharmaceutical composition of the invention. The composition
 CC comprises an effective amount of an ICAM-1 antisense molecule
 CC encapsulated in a lipid mixture, the lipid mixture comprising at
 CC least two members selected from phospholipids, steroids and
 CC cationic lipids. The composition is used in a method to treat
 CC pathological conditions associated with overexpression of ICAM-1, such
 CC as Alzheimer's disease, multiple sclerosis, viral hepatitis,
 CC cholangitis, cardiac allograft rejection, etc.
 SQ Sequence 610 AA;

Query Match 18.9%; Score 298; DB 1; Length 610;
 Best Local Similarity 27.4%; Pred. No. 1,43e-20;
 Matches 45; Conservative 33; Mismatches 79; Indels 7; Gaps 6;
 Db 210 CDRGYLPSSMETMOCSSGEMSA--PIPACNVECDANTNPANGFVECFQNGSPFWMTT 267
 QY 34 CRFGYSLNINIVWCRK-GEWVALNPLRKCQKRCRPGHDPDTFGFTT-LTGNNVFEYGVK 91
 Db 268 CTFDEGEFELMGAQSL-QCTSSGNDNNEKPTCKAVTCRAVROPONGS-VRCSSHAPAGEF 325
 QY 92 AYTTCNEGQYLLGEINRECDTDG-WTNDIPICEVVKCLPVTAPENCKIYSSAMEPDREY 150
 Db 326 TFKSSCNFTCEBGFMLQCPAOCVECTTQOGWTOQIPVCEAFQCTA 369
 QY 151 HFGQAVRFVNCNSGYKIEGDEMHCSDDGFWSKERKPCVEISCKS 194

A portion (W06881) of the complement-inhibitor, decay accelerating factor (DAF), is used in novel chimeric proteins of formula A-R1-B-R-C, where A and C are peptides (W06875-79, W06883-90) able to bind glycosaminoglycans (esp. heparin) present on cell surfaces, CC R1 is a portion of DAF or membrane co-factor protein (MCP, see also W06880), R2 is DAF when R1 is MCP or MCP when R1 is DAF, and B is a peptide that may have complement inhibitor activity. The chimeric proteins (see also W06882) are directed to cell surfaces where they inhibit complement-mediated cell lysis. They are used to treat and prevent disease states in which complement plays a role, e.g. sepsis, adult respiratory distress syndrome, reperfusion injury and tissue damage.

Sequence 299 AA;

Query Match 18.6%; Score 293; DB 1; Length 299;

Best Local Similarity 29.7%; Pred. No. 4,96e-20; Indels 19; Gaps 14;

Matches 58; Conservative 43; Mismatches 75; Indels 19; Gaps 14;

Db 40 C-EVPTRLNSASLKOPYITONTYFVGVVEYECRPGYRREPSLSPKLTCLQNLKMWSTAVE 98

Qy 3 CNELPFRNTETILIGSWSMDQTY-PEGTQAIYKCRPGYRSLGNVT--MVC-RKGEWVALNP 58

Db 99 F--CKKSCNPGEIRNGQIDVPGG-IL-FGATISFCNTGKYLFGSTSSFCILSSSVQ 154

Qy 59 LKCKQKRCPCGHPGDPFTFTLTGWNVEYGVKAVYTCNEGQYLLGELN-YR-ECDDTG- 115

Db 155 WSDPLPECREIYCPAPQIDNG-IIQG--ERDH-YGYRSQVYACNKGFTMIGHSIYCT 210

Qy 116 WTNDIPICEVVKCLPYTAPENKIVSSAMEPDRHYFGQAVRFVNCNSGKIEGDEMHCS 175

Db 211 VNDEGEMSGPPEC 225

Qy 176 --DDGFWSKKPKC 187

RESULT 11

ID W26317 standard; Protein; 381 AA.

AC W26317;

DT 17-NOV-1997 (first entry)

DE Human decay accelerating factor.

KM DAF; decay accelerating factor; human; C3 complement inhibitor;

OS Homo sapiens.

PI Key

FT peptide

FT 1.34

FT /label= sig_peptide

FT /label= Mat_protein

PN US5643770-A.

PD 01-JUL-1997.

PF 21-JUL-1994; 278630.

PR 21-JUL-1994; US-278630.

PA (ALEX-) ALEXION PHARM INC.

PI Mason JM, Squinto SP;

DR WPI: 97-350243/32.

DR N-PSDB: T84471.

PT Retroviral vector particle expressing complement inhibitor activity

PS - for transducing cells in body fluids containing complement

PS Disclosure: Column 41-46; 32pp; English.

CC This protein sequence comprises human decay accelerating factor

CC (DAF), a protein that has C3 complement inhibitor activity.

CC Claimed retroviral vector particles express a complement inhibitor

CC such as DAF, and are thereby protected from inactivation upon

CC exposure to body fluids containing complement. Also claimed are: (1)

CC a producer cell producing the vector particle; and (2) a chimeric

CC retroviral envelope protein (see also W26326) with at least part of

CC the N-terminal receptor-binding domain removed and replaced by a

CC protein domain having a complement inhibitor activity. The vector

CC is used in a claimed method for transducing cells in the presence of

CC a body fluid containing complement, preferably ex vivo, especially

CC for gene therapy, e.g. of hereditary or acquired blood disorders by

CC transduction of hematopoietic stem cells.

CC Sequence 381 AA;

Query Match 18.6%; Score 293; DB 1; Length 381;

Best Local Similarity 29.7%; Pred. No. 4,96e-20; Indels 19; Gaps 14;

Matches 58; Conservative 43; Mismatches 75; Indels 19; Gaps 14;

Db 98 C-EVPTRLNSASLKOPYITONTYFVGVVEYECRPGYRREPSLSPKLTCLQNLKMWSTAVE 156

Qy 3 CNELPFRNTETILIGSWSMDQTY-PEGTQAIYKCRPGYRSLGNVT--MVC-RKGEWVALNP 58

Db 157 F--CKKSCNPGEIRNGQIDVPGG-IL-FGATISFCNTGKYLFGSTSSFCILSSSVQ 212

Qy 59 LKCKQKRCPCGHPGDPFTFTLTGWNVEYGVKAVYTCNEGQYLLGELN-YR-ECDDTG- 115

Db 213 WSDPLPECREIYCPAPQIDNG-IIQG--ERDH-YGYRSQVYACNKGFTMIGHSIYCT 268

Qy 116 WTNDIPICEVVKCLPYTAPENKIVSSAMEPDRHYFGQAVRFVNCNSGKIEGDEMHCS 175

Db 269 VNDEGEMSGPPEC 283

Qy 176 --DDGFWSKKPKC 187

RESULT 12

ID R66683 standard; Protein; 381 AA.

AC R66683;

DT 23-JUL-1995 (first entry)

DE Decay accelerating factor.

KM DAF; mDAF; fusion protein; liposome;

OS cell targeting; glycoposphatidylinositol; GPI; drug delivery.

PI Homo sapiens.

FT key

FT peptide

FT 1.34

FT /label= sig_peptide

FT modified_site

FT /note= "probable phosphatidylinositol

FT derivatization site"

PN US5374548-A.

PD 20-DEC-1994.

PF 02-MAY-1986; 859107.

PR 02-MAY-1986; US-859107.

PR 06-AUG-1987; US-083757.

PR 19-DEC-1991; US-811048.

PR 12-FEB-1993; US-017934.

PA (GETH) GENENTECH INC.

PI Caras IW;

DR WPI: 95-035649/05.

DR N-PSDB: 079863.

PT Liposome(s) for targeting particular cells contg. fusion protein

PT - of glyco:phosphatidyl:inositol anchor and heterologous

PT targeting protein, e.g. for delivering toxins to infected or

PT cancer cells.

PS Disclosure: Fig. 1a-1f; 36pp; English.

CC A probe (given in Q79865) based on the N-terminal sequence of human

CC decay accelerating factor (DAF) was used to screen a HeLa cell

CC lambda cDNA library. Isolated clones encoding membrane-bound DAF

CC (mDAF) were obtained, and the full sequence of mDAF cDNA was

CC determined (Q79863). The GPI signal domain of mDAF may be fused to

CC a heterologous protein and targeted to cell membrane surfaces.

CC Sequence 381 AA;

CC Query Match 18.6%; Score 293; DB 1; Length 381;

CC Best Local Similarity 29.7%; Pred. No. 4,96e-20; Indels 19; Gaps 14;

CC Matches 58; Conservative 43; Mismatches 75; Indels 19; Gaps 14;

Db 98 C-EVPTRLNSASLKOPYITONTYFVGVVEYECRPGYRREPSLSPKLTCLQNLKMWSTAVE 156

Qy 3 CNELPFRNTETILIGSWSMDQTY-PEGTQAIYKCRPGYRSLGNVT--MVC-RKGEWVALNP 58

Db 157 F--CKKSCNPGEIRNGQIDVPGG-IL-FGATISFCNTGKYLFGSTSSFCILSSSVQ 212

Qy 59 LKCKQKRCPCGHPGDPFTFTLTGWNVEYGVKAVYTCNEGQYLLGELN-YR-ECDDTG- 115

Db 213 WSDPLPECREIYCPAPQIDNG-IIQG--ERDH-YGYRSQVYACNKGFTMIGHSIYCT 268

OY 116 WTNDIPICEVVKCLPVTAPENGKIVSSAMEPDRREYHFGQAVRVCNSGYRIEEDDEMHCS 175
 DB 269 VNNDGEMSGPPEPC 283
 OY 176 ---DDGFWSKEKPKC 187

RESULT 13
 ID P70048 standard; protein; 381 AA.
 AC P70048:
 DT 03-FEB-1991 (first entry)
 DE Human decay acceleration factor variant #1.
 KM Decay acceleration factor.
 FH Key Location/Qualifiers
 FT region 331..347
 PI /label-putative transmembrane region
 PD EP-244267-A.
 PF 04-NOV-1987.
 PR 01-MAY-1987; 303944.
 PA 02-MAY-1986; US-859107.
 PI (GERTH) Genentech Inc.
 DR Caras IW;
 DR WPI: 87-308481/44.
 DR N-PSDB: N70047.
 PT New decay accelerating factor variants - obtained with the factor by
 using recombinant DNA procedures.
 PS Disclosure; Page 15-17; 20pp; English.
 CC The probable phosphatidylinositol derivatisation site is Cys(330).
 CC The DAF variant is useful for treating paroxysmal nocturnal
 CC haemoglobinuria, or inflammatory or cell lytic autoimmune
 CC diseases. It may be used to ameliorate allograft rejection
 CC or autoimmune diseases. See also N70046, N70048.
 SQ Sequence 381 AA;

Query Match 18.6%; Score 293; DB 1; Length 381;
 Best Local Similarity 29.7%; Pred. No. 4,96e-20;

Matches 58; Conservative 43; Mismatches 75; Indels 19; Gaps 14;

DB 98 C-EVPTRLNSASLKQPTITQNTFPVGTVEYECRPGYRREPSLSKTLCTQLNKLKSTAVE 156
 OY 3 CNELPRRNTTEILTGSWSDQTY-PEGTQAIYKCRPGYRSLGNVI--MVC-RKGEWVALNP 58
 DB 157 F--CKKSCPNNGEIRNGQIDVPGG-IL-FGATISFSCNTGKTLGTSSTFCLISGSSVQ 212
 OY 59 LKCKQKRCGHPGDTPTGFTLTGNGVFEYGVKAVYTCNEGTLGELIN-YR-ECDTDG- 115
 DB 213 WSDPLPECREIYCAPPQIDNG-IIQG--ERDH-YGYROSVTYACNKGFTMIGESHIYCT 268
 OY 116 WTNDIPICEVVKCLPVTAPENGKIVSSAMEPDRREYHFGQAVRVCNSGYRIEEDDEMHCS 175
 DB 269 VNNDGEMSGPPEPC 283
 OY 176 ---DDGFWSKEKPKC 187

RESULT 14
 ID P94773 standard; protein; 381 AA.
 AC P94773:
 DT 04-JUL-1990 (first entry)
 DE Decay accelerating factor (DAF) of clones lambda 33 and lambda 47.
 KM DAF; allograft rejection; affinity purification;
 KM autoimmune disease; ds.
 OS Synthetic.
 PN W08901041-A.
 PD 09-FEB-1989.
 PR 3-AUG-1988; 02648.
 PR 6-AUG-1987; US-083757.
 PA (GERTH) Genentech Inc.
 PI Caras I;
 DR WPI: 89-061177/08.
 DR N-PSDB: N91043.
 PT Fusion polypeptide for targeting protein to cell membrane -

PT comprises phospholipid anchor domain with heterologous
 PT polypeptide.
 PS Disclosure; 61pp; English.
 CC Recombinant DAF's are useful in treatment of inflammatory or cell lytic
 CC autoimmune diseases and allograft rejection. Useful in diagnostic
 CC compositions or in affinity purification.
 SQ Sequence 381 AA;

Query Match 18.6%; Score 293; DB 1; Length 381;
 Best Local Similarity 29.7%; Pred. No. 4,96e-20;

Matches 58; Conservative 43; Mismatches 75; Indels 19; Gaps 14;

DB 98 C-EVPTRLNSASLKQPTITQNTFPVGTVEYECRPGYRREPSLSKTLCTQLNKLKSTAVE 156
 OY 3 CNELPRRNTTEILTGSWSDQTY-PEGTQAIYKCRPGYRSLGNVI--MVC-RKGEWVALNP 58
 DB 157 F--CKKSCPNNGEIRNGQIDVPGG-IL-FGATISFSCNTGKTLGTSSTFCLISGSSVQ 212
 OY 59 LKCKQKRCGHPGDTPTGFTLTGNGVFEYGVKAVYTCNEGTLGELIN-YR-ECDTDG- 115
 DB 213 WSDPLPECREIYCAPPQIDNG-IIQG--ERDH-YGYROSVTYACNKGFTMIGESHIYCT 268
 OY 116 WTNDIPICEVVKCLPVTAPENGKIVSSAMEPDRREYHFGQAVRVCNSGYRIEEDDEMHCS 175
 DB 269 VNNDGEMSGPPEPC 283
 OY 176 ---DDGFWSKEKPKC 187

RESULT 15
 ID W73505 standard; protein; 381 AA.
 AC W73505:
 DT 01-MAR-1999 (first entry)
 DE Decay accelerating factor protein.
 KM DAF; decay accelerating factor; human; complement protein; gene therapy;
 KM viral vector; ds.
 OS Homo sapiens.
 PN J10313865-A.
 PD 02-DEC-1998.
 PR 15-MAY-1997; 125965.
 PR (DINA-) DINABEKRU KENKYUSHO KK.
 DR WPI: 99-074147/07.
 DR N-PSDB: V08935.
 PT Vector having complement controlling factor - useful for gene
 PT therapy.
 PS Example 3; Page 8-9; 15pp; Japanese.
 CC This sequence is the human decay accelerating factor (DAF) protein.
 CC DAF can be used in the viral vector of the invention. The viral vector
 CC contains a factor controlling the function of a human complement protein,
 CC particularly a membrane combining type protein. The viral vector, which
 CC is stable in situ, is useful for gene therapy.
 SQ Sequence 381 AA;

Query Match 18.6%; Score 293; DB 1; Length 381;
 Best Local Similarity 29.7%; Pred. No. 4,96e-20;

Matches 58; Conservative 43; Mismatches 75; Indels 19; Gaps 14;

DB 98 C-EVPTRLNSASLKQPTITQNTFPVGTVEYECRPGYRREPSLSKTLCTQLNKLKSTAVE 156
 OY 3 CNELPRRNTTEILTGSWSDQTY-PEGTQAIYKCRPGYRSLGNVI--MVC-RKGEWVALNP 58
 DB 157 F--CKKSCPNNGEIRNGQIDVPGG-IL-FGATISFSCNTGKTLGTSSTFCLISGSSVQ 212
 OY 59 LKCKQKRCGHPGDTPTGFTLTGNGVFEYGVKAVYTCNEGTLGELIN-YR-ECDTDG- 115
 DB 213 WSDPLPECREIYCAPPQIDNG-IIQG--ERDH-YGYROSVTYACNKGFTMIGESHIYCT 268
 OY 116 WTNDIPICEVVKCLPVTAPENGKIVSSAMEPDRREYHFGQAVRVCNSGYRIEEDDEMHCS 175
 DB 269 VNNDGEMSGPPEPC 283
 OY 176 ---DDGFWSKEKPKC 187

Fri Jun 9 10:53:44 2000

US-09-316-163-9.rag

Page 7

Search completed: Thu Jun 8 21:37:19 2000
Job time : 11 secs.

 W39154
 (TW)

Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

Msearch_p protein - protein database search, using Smith-Waterman algorithm
 Run on: Thu Jun 8 21:40:50 2000; Maspar time 9.79 Seconds
 641.132 Million cell updates/sec
 Tabular output not generated.

Title: >US-09-316-163-10
 Description: (1-265) from US09316163.pep
 Perfect Score: 2029
 Sequence: 1 EDCEMLPPRRMTILITGSWS.....EKSCDNPYIPNGDYSPLRIK 265

Scoring table: PAM 150
 Gap 11

Searched: 188963 seqs/23686106 residues

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database: a:geneseq36
 1:geneseqp

Statistics: Mean 32.240; Variance 119.188; scale 0.271

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	1429	70.4	240	1	Human partial Complement	1.09e-147
2	1298	64.0	216	1	Clone PRB89FH410 CFH r	1.28e-132
3	382	18.8	581	1	Human C4 binding prote	4.34e-29
4	371	18.3	263	1	Deduced sequence of co	6.80e-28
5	344	17.0	830	1	P-selectin.	5.67e-25
6	341	16.8	302	1	Herpesvirus secreted g	1.19e-24
7	341	16.8	302	1	Herpesvirus salm1r1 SC	1.19e-24
8	341	16.8	360	1	Herpesvirus salm1r1 g	1.19e-24
9	341	16.8	360	1	Herpesvirus salm1r1 mc	1.19e-24
10	335	16.5	376	1	CD46 wild-type.	5.29e-24
11	335	16.5	377	1	CD46 construct subSC33	5.29e-24
12	335	16.5	377	1	Human MCP.	5.29e-24
13	335	16.5	384	1	Human membrane cofacto	5.29e-24
14	335	16.5	384	1	Human CD46.	5.29e-24
15	333	16.4	279	1	Human membrane cofacto	8.68e-24
16	333	16.4	279	1	Human membrane cofacto	8.68e-24
17	332	16.4	324	1	CD46 from clone pm5.8.	1.11e-23
18	332	16.4	324	1	CD46 from clone pm5.8.	1.11e-23
19	332	16.4	324	1	CD46 from clone pm5.1.	1.11e-23
20	330	16.3	251	1	Membrane co-factor prot	1.82e-23
21	330	16.3	254	1	Membrane co-factor pro	1.82e-23
22	330	16.3	357	1	Human membrane cofacto	1.82e-23
23	330	16.3	577	1	Membrane co-factor pro	1.82e-23

RESULT ID	Score	Query Match	Length	ID	Description	Pred. No.
1	1429	70.4	240	1	Human partial Complement	1.09e-147
2	1298	64.0	216	1	Clone PRB89FH410 CFH r	1.28e-132
3	382	18.8	581	1	Human C4 binding prote	4.34e-29
4	371	18.3	263	1	Deduced sequence of co	6.80e-28
5	344	17.0	830	1	P-selectin.	5.67e-25
6	341	16.8	302	1	Herpesvirus secreted g	1.19e-24
7	341	16.8	302	1	Herpesvirus salm1r1 SC	1.19e-24
8	341	16.8	360	1	Herpesvirus salm1r1 g	1.19e-24
9	341	16.8	360	1	Herpesvirus salm1r1 mc	1.19e-24
10	335	16.5	376	1	CD46 wild-type.	5.29e-24
11	335	16.5	377	1	CD46 construct subSC33	5.29e-24
12	335	16.5	377	1	Human MCP.	5.29e-24
13	335	16.5	384	1	Human membrane cofacto	5.29e-24
14	335	16.5	384	1	Human CD46.	5.29e-24
15	333	16.4	279	1	Human membrane cofacto	8.68e-24
16	333	16.4	279	1	Human membrane cofacto	8.68e-24
17	332	16.4	324	1	CD46 from clone pm5.8.	1.11e-23
18	332	16.4	324	1	CD46 from clone pm5.8.	1.11e-23
19	332	16.4	324	1	CD46 from clone pm5.1.	1.11e-23
20	330	16.3	251	1	Membrane co-factor prot	1.82e-23
21	330	16.3	254	1	Membrane co-factor pro	1.82e-23
22	330	16.3	357	1	Human membrane cofacto	1.82e-23
23	330	16.3	577	1	Membrane co-factor pro	1.82e-23

ALIGNMENTS

RESULT 1
 ID W39154 standard; Protein; 240 AA.
 AC W39154;
 DT 27-APR-1998 (first entry)
 DE Human partial Complement factor H protein fragment 1.
 KW Complement factor H; tumour associated antigen; renal cancer;
 KW urogenital cancer; medicament; modulator.
 OS Homo sapiens.
 PN WO9738136-A1.
 PD 16-OCT-1997.
 PE 09-APR-1997; U05710.
 PR 06-MAR-1997; US-812481.
 PR 09-APR-1996; US-015083.
 PR 09-APR-1996; US-630048.
 PR 06-MAR-1997; US-038614.
 PA (BARD-) BARD DIAGNOSTIC SCI INC.
 PI Enfield DL, Hass GM, Kinders RJ;
 PI WPI: 97-512742/47.
 DR N-PSDB: V02790.
 PT Treating or screening for cancer, e.g. renal or urogenital cancer -
 PT by modulating or detecting tumour associated human complement factor
 PT H related antigen, or nucleic acid encoding it
 PS Example 6B: Fig 6B: 104pp; English.
 CC This partial protein sequence represents a region of the human
 CC tumour-associated complement factor H (CFH). This sequence is used
 CC in the identification of complement factor H related proteins and
 CC such proteins and a CFH antigens can be used in screening for the
 CC treatment of renal or urogenital cancer, e.g. bladder, cervical or
 CC prostate cancer. Agents that may modulate this antigen could be used in
 CC the manufacture of a medicament for the treatment of a tumour cell.
 SQ Sequence 240 AA:
 Query Match 70.4%; Score 1429; DB 1; Length 240;
 Best Local Similarity 100.0%; Pred. No. 1.09e-147;
 Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 FTLLGVNVEFGVAVYTCNKGYYLLGEINVRRCDDTDGWTNDIPICVVKCLPVTAPENG 60
 78 FTLLGVNVEFGVAVYTCNKGYYLLGEINVRRCDDTDGWTNDIPICVVKCLPVTAPENG 137
 Oy 61 KIVSSAMEPDRHYFGAVRVCNSGYKIGDEMHCSDDGFSKPKVETISCSPPV 120
 Oy 138 KIVSSAMEPDRHYFGAVRVCNSGYKIGDEMHCSDDGFSKPKVETISCSPPV 197
 Db 121 INGSISGKIITKNEFRVYCNMGYSERGDVAVCESGMRPLPSCSEKSCNPPIPN 180

QY 198 INGSPIKIKYKENERFOYKCMNGYERSDAVCTESGMRPLPSCSEKSCDNPIYING 257
 Db 181 DYSPLRIK 188
 QY 258 DYSPLRIK 265

RESULT 2
 ID W39155 standard; Protein; 216 AA.

AC W39155;
 DT 27-APR-1998 (first entry)
 DE Clone PRB9FH410 CFH related protein fragment.
 KW Complement factor H; tumour associated antigen; renal cancer;
 KW urogenital cancer; medicament; modulator.
 OS Synthetic.
 PN WO9738136-A1.
 PD 16-OCT-1997.
 PF 09-APR-1997; U05710.
 PR 06-MAR-1997; US-812481.
 PR 09-APR-1996; US-015083.
 PR 06-MAR-1997; US-038614.
 PR 06-MAR-1996; US-630048.
 PA (BARD-) BARD DIAGNOSTIC SCI INC.
 PI Enfield DL, Hass GM, Kinders RJ;
 DR WPI: 97-512742/47.
 DR N-PSDB: V02791.

PT Treating or screening for cancer, e.g. renal or urogenital cancer -
 PT by modulating or detecting tumour associated human complement factor -
 PT H related antigen, or nucleic acid encoding it
 PS Example 6B; Fig 6B; 104pp; English.
 CC This partial protein is found in clone PRB9FH410 and represents a
 CC complement factor H related protein with homology to a region of the
 CC human tumour-associated complement factor H (CFH). The detection of this
 CC protein and a CFH antigen can be used in screening or for the treatment
 CC of renal or urogenital cancer, e.g. bladder, cervical or prostate cancer.
 CC Agents that may modulate this antigen could be used in the manufacture of
 CC a medicament for the treatment of a tumour cell.
 SQ Sequence 216 AA;

Query Match 64.0%; Score 1298; DB 1; Length 216;
 Best Local Similarity 99.4%; Pred. No. 1,28e-132;
 Matches 170; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 1 TCNEGYQLLEINRECDTGDWTNDIPICFVVKCLPYTAPENKITYSSAMEPDRHFQ 60
 QY 95 TCNEGYQLLEINRECDTGDWTNDIPICFVVKCLPYTAPENKITYSSAMEPDRHFQ 154
 Db 61 AVRFVCSGKIRIGDEMHGSDGFMGKPKCVCVEISCKSPDIVINGSPISOKIYKENER 120
 QY 155 AVRFVCSGKIRIGDEMHGSDGFMGKPKCVCVEISCKSPDIVINGSPISOKIYKENER 214
 Db 121 FOYKCMNGYERSDAVCTESGMRPLPSCSEKSCDNPIYINGDYSPLRIK 171
 QY 215 FOYKCMNGYERSDAVCTESGMRPLPSCSEKSCDNPIYINGDYSPLRIK 265

RESULT 3
 ID R13490 standard; Protein; 581 AA.
 AC R13490;
 DT 30-OCT-1991 (first entry)
 DE Human C4 binding protein.
 KW C4bp; monomer; complement protein; pI 0.04; C4bp.3; SCR;
 KW short consensus repeat.
 OS Homo sapiens.
 FH Key
 FT peptide
 FT Location/Qualifiers
 FT 1..32
 FT /label= signal_peptide
 FT protein
 FT 33..581
 FT /label= C4bp
 FT region
 FT 33..93
 FT /label= SCR8
 FT 94..155
 FT /label= SCR7

FT region
 FT 156..219
 FT /label= SCR6
 FT region
 FT 220..279
 FT /label= SCR5
 FT region
 FT 280..345
 FT /label= SCR4
 FT region
 FT 346..406
 FT /label= SCR3
 FT region
 FT 407..464
 FT /label= SCR2
 FT region
 FT 465..523
 FT /label= SCR1
 FT domain
 FT 524..581
 FT /label= C4bp_core
 FT /note= "responsible for multimer assembly"
 FT disulfide_bond
 FT 34..80
 FT /note= "intradomain"
 FT disulfide_bond
 FT 65..92
 FT /note= "intradomain"
 PN MO9111461-A.
 PD 08-AUG-1991.
 PF 28-JAN-1991; U00567.
 PR 26-JAN-1990; US-470888.
 PA (BIOG-) BIOGEN INC.
 PI Pasak MP, Winkler G, Liu TR;
 DR WPI: 91-252613/34.
 DR N-PSDB: Q13242.
 PT New C4 binding protein fusion proteins and DNA encoding them -
 PT comprise assemblies of C4bp monomers linked to functional moiety,
 PT e.g. A2T, useful as delivery vehicles in diagnosis and therapy
 PS Example 1; Fig 1; 105pp; English.
 CC This sequence was deduced from human hepatocyte (Hep G2) cDNA
 CC obtained following PCR amplification. The protein is a monomer
 CC containing 8 SCRs. Each SCR forms a looped domain due to the
 CC presence of two intradomain disulfide bonds (only the disulfide
 CC bonds of SCR8 are labelled in the Features Table). Within each SCR,
 CC the first cysteine residue bonds with the third and the second
 CC cysteine residue bonds with the fourth. This secondary structure is
 CC responsible for the conformational flexibility of the C4bp monomer.
 CC The invention covers fusion proteins in which the monomer sequence,
 CC or a specified subfragment of it, e.g. having 5', 4', 3 or 1 SCR(s)
 CC is fused to the C-terminal of a protein such as a viral receptor,
 CC cell ligand, a bacterial, viral or parasitic immunogen, enzyme,
 CC cytokine, toxin, etc. See also Q13243-51.
 SQ Sequence 581 AA;

Query Match 18.8%; Score 382; DB 1; Length 581;
 Best Local Similarity 31.5%; Pred. No. 4.34e-29;
 Matches 79; Conservative 52; Mismatches 98; Indels 22; Gaps 20;

Db 51 TETRFGTTLTKYCLPGYRSHSTQTLTNSDGEWY-VYTFCTY-KR-CRHPGELRNQ 107
 QY 20 SDQTYPEGTQAIYKCRGY-RSLGNVIMCRK-GEWALNPLKCKCRGPHGDDPFET 77
 Db 108 VEIKT-DL-SFGSIEFSCSEGFLLIGSTTSKCEVODRGVMSHPLPOCEIYKCKPPI 165
 QY 78 FTLLGAVVEFYGVKAVYTCNEGYQLLEINR-EC-DTD-GWTNDIPICFVVKCLPYTAP 134
 Db 166 RNGR--HSG-E-ENFYAYGGSVYSCDPRSLIGHASISCTYENETIYWRSPPCETI 221
 QY 135 ENGRIVSSAMEPDRHFQAVRFVCSGKIRIGDEMHGSDGFMGKPKCVCVEISCKSPDIVINGSPISOKIYKENER 190
 Db 222 TCRKPDVSHGEMVSGFPIYKDTIVFCCKGFGVLRGSSVYHCDADSFWNPSPACEPN 281
 QY 191 SCKSPDIVINGSPISOKIYKENERFOYKCMNGYERSDAVCTESGMRPLPSCSEKSCDNPIYING 247
 Db 282 SCINLPDIIPA 292
 QY 248 SCIN-PIYING 257

RESULT 4
 ID P92003 standard; Protein; 263 AA

CC	P92003	09-FEB-1990	(first entry)
DR	DE	Deduced sequence of complement 4b (C4b) binding protein	
KM	KM	Vaccinia virus WR strain: anti-complement protein; complement	
FM	KM	inactivators; complement 4b; C4b.	
FT	FT	Key	Location/Qualifiers
FT	FT	peptide	1..19
FT	FT	duplication	/label= signal_peptide
FT	FT	duplication	20
FT	FT	duplication	/note= "See note a in comments below."
FT	FT	duplication	83
FT	FT	duplication	/note= "See note a in comments below."
FT	FT	duplication	144
FT	FT	duplication	/note= "See note a in comments below."
FT	FT	duplication	202
FT	FT	duplication	/note= "See note a in comments below."
PN	PN	US-7239208-A.	
PD	PD	14-MAR-1989.	
PE	PE	20-AUG-1988; 239208.	
PR	PR	20-AUG-1988; US-239208.	
PA	PA	(USSH) Nat Inst of Health.	
PI	PI	Korval G:	
DR	DR	WPI; 89-165451/22.	
PT	PT	N-PSDB; N90113.	
PT	PT	New protein with anti-complement activity	
PS	PS	- encoded by Vaccinia virus 35k gene	
CC	CC	Disclosure; Figure 2A; 20pp; English.	
CC	CC	C4b-binding protein which specifically blocks human complement cascades.	
CC	CC	It is the deduced sequence of a 35kDa protein encoded by sequence 52-8404	
CC	CC	of the 35k gene of vaccinia virus strain WR. Note a - these sites	
CC	CC	indicate the start of 60 amino acid tandem repeating units which have a	
CC	CC	consensus sequence. The signal peptide sequence is not found in purified	
CC	CC	35k protein recovered from the medium of cells infected with vaccinia	
CC	CC	virus strain WR. A suggested use is to treat diseases due to abnormally	
CC	CC	high complement activity.	
CC	CC	Sequence 263 AA:	

Query Match	18.3%;	Score 371;	DB 1;	Length 263;
Best Local Similarity	33.8%;	Pred. No. 6.80e-28;		
Matches	79;	Conservative	37;	Mismatches 97;
			Indels	21;
			Gaps	17;

```

Db      40  ANNANVIGITIELLPYGRKQMGPIYAKGTGWM-FL--FNQCIKRRCSPFDINDGQ 96
Qy      20  SDQTYIEGQAIYKRCPRGRS--LGNVIMWCKRGEMVALNLRKQCRPRGHDPEDTFEFT 77
Db      97  LDI-GG-V-DFSSSTITYSCNSGYHLIGESKSYCELGSTGSMWNNDPEPISRYVACQSPFS 153
Qy      78  FTLTGGNVEYEGYKVAVYTCNEGYYOLGLEIN-YRE-CDPDG--WNNDPIICEVVKCLEPTA 133
Db      154  ISNGR--HNGYE-DE-YTGTGAVYTCNSNGYSLIENSGVLCSGGE-WG-DEPTQIYKCP 207
Qy      134  PENGKIVSSAMMEDRHYHGGAVRFVNCNGYLIIEDEEMHCSDDGFWMKRKPVEISELK 193
Db      208  HPTISNGYLSGFKRSYTYNDVDEKCKYKGYKLSSSSSTGSPGNTMKPELPKC 261
Qy      194  SPDVINGSPISQ-KIYKENERFOYKCMNGYERYSGDVACTESG-WRP-LPSC 244

RESULT      5
ID      R65216 standard; Protein: 830 AA.
AC      R65216;
DT      04-OCT-1995 (first entry)
DE      P-selectin.
KW      Control elements; gene expression; endothelial cells;
        megakaryocytes; rheumatoid arthritis; Ischaemic injury;
        kw atherosclerosis; bacterial sepsis; tumour metastases.
OS      Homo sapiens.
PN      M09506118-A.
PD      02-MAR-1995.
PF      19-AUG-1994; U09395.
PR      20-AUG-1993; U5-110158.
PI      (OKLA ) UNIV OKLAHOMA STATE.
        Mcever RP, Pan J;

```

DR WPI: 95-106847/14:
PT Expression control/1:
PT P-selectin gene - for regulating gene expression in endothelial
PT cells and mega:karyocyte(s)
PS Disclousure; Fig 1; 88pp; English.
CC The sequence is that of the P-selectin gene product from endothelial
CC cells, determined from a composite of four overlapping cDNAs: lambda
CC GMEB1, lambda GMEB4. The P-selectin gene or its fragments may be used
CC to inhibit P-selectin expression, thus controlling inflammatory and
CC haemostatic processes in e.g. rheumatoid arthritis, ischaemic injury,
CC bacterial sepsis, tumour metastases, atherosclerosis, etc.
SQ Sequence 830 AA;

Query Match	17.0%;	Score 344;	DB 1;	Length 830;
Best Local Similarity	26.0%;	Pred. No. 5.67e-25;		
Matches	69;	Conservative	56;	Mismatches 122;
			Indels	18;
			Gaps	18

[illegible]

RESULT	6
ID	W26320 standard; Protein; 302 AA.

DT	17-NOV-1997	(first entry)
DE	Herpesvirus secreted glycoprotein SCGPB.	
KW	CCPB; C3 complement inhibitor; retrovirus; vector; gene therapy;	
OS	stem cell.	
FH	Herpesvirus Saimiri strain #1.	
FT	Key	
FT	Location/Qualifiers	
FT	1..20	
FT	/label= Sig_peptide	
FT	21..360	
FT	/label= Mat_protein	
PN	US5643770-A.	
PD	01-JUL-1997.	
PF	21-JUL-1994; 278630.	
PA	21-JUL-1994; US-278630.	
PI	(ALEX-) ALEXION PHARM INC.	
PI	Mason JM, Squinto SP;	
DR	WPI: 97-350243/32.	
DR	N-PSDB: T84474.	
PT	Retroviral vector particle expressing complement inhibitor activity	
PT	- for transducing cells in body fluids containing complement	
PS	Disclosure; Column 53-56; 32pp; English.	
CC	This protein sequence comprises the herpesvirus Saimiri secreted	
CC	glycoprotein SCGPB, an inhibitor of complement C3. Claimed	
CC	retroviral vector particles express a complement inhibitor such as	
CC	CCPB, and are thus protected from inactivation upon exposure to	
CC	body fluids containing complement. Also claimed are: (1) a	
CC	producer cell producing the vector particle; and (2) a chimeric	
CC	retroviral envelope protein (see also W26324) with at least part	
CC	of the N-terminal receptor-binding domain removed and replaced by	
CC	a protein domain having a complement inhibitor activity. The	
CC	vector is used in a claimed method for transducing cells in the	
CC	presence of a body fluid containing complement, preferably ex vivo,	

KM MCPPH; membrane complement control protein homolog; SCCPH;
 KM secreted complement control protein homolog; HVS-15;
 KM complement regulatory protein; complement inhibitory protein.
 OS Herpesvirus saimiri.
 FT Key Location/Qualifiers
 FT 1-20
 FT /label= Sig-peptide
 PN WO9416062-A.
 PD 21-JUL-1994.
 PF 12-JAN-1993; U00672.
 PR 12-JAN-1993; WO-U00672.
 PA (ALEX-) ALEXION PHARM INC.
 PI Albrecht J. Fleckenstein B;
 DR WPI; 94-249210/30.
 DR N-PSDB; Q66956.
 PT New complement regulatory proteins of Herpes virus Saimiri - are
 PT used to inhibit complement-mediated lysis in treating immune
 PT disorders and to develop other prods.
 PS Disclosure: Page 33-35; 59pp; English.
 CC Gene sequences for 3 complement regulatory proteins encoded within
 CC the genome of herpesvirus saimiri (HVS) are disclosed, i.e. membrane
 CC complement control protein homolog (mccph) (Q66956), secreted
 CC complement control protein homolog (sccph) (Q66957) and HVS-15
 CC (Q66958). MCPPH and SCCPH share substantial homology with the human
 CC complement inhibitory proteins factor H, CD35, CD46, CD55 and C4bp,
 CC which inhibit C3 convertase activity in the complement cascade. The
 CC gene sequences and corresponding proteins (R5792-94) can be used to
 CC control the complement arm of the immune system.
 SQ Sequence 360 AA;
 Query Match 16.8%; Score 341; DB 1; Length 360;
 Best Local Similarity 31.9%; Pred. No. 1.19e-24;
 Matches 74; Conservative 45; Mismatches 91; Indels 22; Gaps 17;
 Db 44 YPNGTTHVTCREBYAKRPVOTVCVGNW-TV-P-KKCKKKKSTPODLLNGRYTGTG- 99
 Qy 24 YPESTQAIYKCRPGYRSIGVYKRGGEVWALNPKRCORPCGHHGDPFGFTLTGG 83
 Db 100 NLV-YGSVITTCNSGYSLIGST-SACLLKRGVWTPRPICDJKKCP--POIAN 155
 Qy 84 NFEYGVKAVTCNEGTLLEINREC--DTDG---WTNDIPICEVYKCLPAPVAPENGK 138
 Db 156 GTHNVK-DF-YTYLDYTVSCNDETKLTLPSPSKLCSSETGSVWNGETKCEIFCKLP 213
 Qy 139 IVSSAMEDEHYHFGQAVFVNCNSGYKIE--GDEMHCSDDGFV-SKRPCKVCVSIKSP 195
 Db 214 QVANAIVYKRSATSMOYLHNVKCYKFMXGEPNT-CHHGWSAIPPC 264
 Qy 196 DVING-SPISOKIITYKNERFOYKCMNGE-YSEKGDVAVCTESGWRP-LPSC 244
 RESULT 10
 ID R93939 standard; Protein; 376 AA.
 AC R93939;
 DT 20-MAY-1996 (first entry)
 DE CD46 wild-type.
 KM CD46; recombinant protein; short consensus repeat; SCR;
 KM regulator of complement activation; transgenic animal; pig;
 KM organ transplantation.
 OS Homo sapiens.
 OS Home sapiens.
 PN WO9606937-A1.
 PD 07-MAR-1996.
 PF 30-AUG-1995; AU0553.
 PR 30-AUG-1994; AU-007724.
 PA (AUST-) AUSTIN RES INST.
 PI Christiansen D. Loveland B. McKenzie IFC, Milland J;
 DR WPI; 96-160368/16.
 DR N-PSDB; T17597.
 PT Increasing prodn. of recombinant proteins, esp. CD46 - by reducing
 PT the amt. of A and/or T in an A and/or T rich region of encoding gene
 PT exon
 PS Disclosure: Page 32-33; 60pp; English.
 CC A cDNA sequence (T17595) codes for wild-type CD46 (R93939), a

CC regulator of complement activation. Expression of CD46 in
 CC eukaryotic host cells, e.g. COS-7 and CHO-K1, is facilitated
 CC by lowering the A+T content of the short consensus repeat regions
 CC of the CD46 gene (see T17596-99) to give products (R93940-43)
 CC useful e.g. as immunomodulators. Modified CD46 may also be
 CC expressed on the organs of transgenic animals of use for organ
 CC transplantation.
 SQ Sequence 376 AA;
 Query Match 16.5%; Score 335; DB 1; Length 376;
 Best Local Similarity 31.0%; Pred. No. 5.29e-24;
 Matches 81; Conservative 54; Mismatches 99; Indels 27; Gaps 22;
 Db 32 DACEE-PTPEFAMELICK-PPRYEIGERVDYKCKKGYFIIPLATITCDRHHMPLVPS 89
 Qy 1 EDCNLLPFRNRTEITLGSWSQDTPEGTQAIYKCRPGYRSIGVY-I-MVC-RKGEWALN 57
 Db 90 D-DACRETCTPIRD-PLNGQAVPANGTYERGQMHFICNGYLLIGELICLGSVA 147
 Qy 58 PLRKQKRPCCGHPGDPPTGFTLTGAVNEFYGVKAVYICNGYLLGE-IVYRE-CDTDG 115
 Db 148 IMSGKPICEKYLCTPPPKIKKHTFSEVEV-FEY-L-DAVYSCDPAPGDPFSLIGE 204
 Qy 116 -WTNDIPICEVYKCLPVAPENGKIVSSAMPDREYHFGAVRVCN-S-G---YKIEGD 169
 Db 205 STTYGDSVWSRAAPDECKVYKCRPVYENGKQISGFEKKRYATWFE--CDKGF-YL 261
 Qy 170 EEMHSDGDFWSEKRPCKVEISKSPDVINGSPIS--OKIITYENERFOYKCMNGEYS 226
 Db 262 DGSDFVYCDNSNTWDPVPVKC 282
 Qy 227 ERGDA-VC-TESGV-RPLPSC 244
 RESULT 11
 ID R93941 standard; Protein; 377 AA.
 AC R93941;
 DT 21-MAY-1996 (first entry)
 DE CD46 construct subSCR3.
 KM CD46; recombinant protein; short consensus repeat; SCR;
 KM regulator of complement activation; transgenic animal; pig;
 KM organ transplantation.
 OS Synthetic.
 PN WO9606937-A1.
 PD 07-MAR-1996.
 PF 30-AUG-1995; AU0553.
 PR 30-AUG-1994; AU-007724.
 PA (AUST-) AUSTIN RES INST.
 PI Christiansen D. Loveland B. McKenzie IFC, Milland J;
 DR WPI; 96-160368/16.
 DR N-PSDB; T17597.
 PT Increasing prodn. of recombinant proteins, esp. CD46 - by reducing
 PT the amt. of A and/or T in an A and/or T rich region of encoding gene
 PT exon
 PS Claim 12; Page 36-37; 60pp; English.
 CC CD46 subSCR3 (R93941) is the product of a cDNA construct
 CC (T17597) chd. by splice overlap extension PCR of wild-type CD46
 CC cDNA (T17595). The A+T content of AT-rich exon 5 of the CD46
 CC gene, encoding the short consensus repeat 3 (SCR3) region of
 CC CD46, is lowered. This results in improved prodn. in eukaryotic
 CC host cells, e.g. CHO-K1 and COS-7. CD46 subSCR3 is used to prevent
 CC complement- or inflammation-mediated tissue damage, to improve
 CC immunity to tumours or viruses, to control fertilisation and to
 CC prevent spontaneous abortion. Expression in transgenic animals,
 CC esp. pigs, provides organs suitable for transplantation.
 SQ Sequence 377 AA;
 Query Match 16.5%; Score 335; DB 1; Length 377;
 Best Local Similarity 31.0%; Pred. No. 5.29e-24;
 Matches 81; Conservative 54; Mismatches 99; Indels 27; Gaps 22;
 Db 33 DACEE-PTPEFAMELICK-PPRYEIGERVDYKCKKGYFIIPLATITCDRHHMPLVPS 90

QY 1 EDCNEIPRRNTEILGNSWSDQTYPEGTQAIKCRPGYSLGNV-I-MVC-RKGEWALN 57
 Db 91 D-DACYRETCPYIRD-PLNGQAVPANGTYEFGYOMHFCINEGYLLIGETILCELKGSVA 148
 QY 58 PLKCCQKRCRPGHGDTPFGFTLLGNGVFEYGVKAVYTNEGQYLLIGE-INYRE-CDTGG 115
 Db 149 IMSGRPICKEKYLCTPPPKIKNGKHTFSEYEV-FEY-L-DAVYSCDPAAGPDPFSLIGE 205
 QY 116 -WTNDIPICEVVKCLPVTAPENCKIYSSAMEPDRHYHFGQAVFYCN-S-G---YKIEGD 169
 Db 206 STIYCGNSVWSRAAPCEKRVKCRFPVYENGKQISGFGKKFYKATVME--CDKGF-YL 262
 QY 170 EEMHCSDGDFWMSKEKRCVCEISCKSPDVINGSPIS---QKIYKENERFYQKCNMGYEYS 226
 Db 263 DGSDDIVCDNSNTWDPVPVKC 283
 QY 227 ERGDA-VC-TESGM-RPLPSC 244

RESULT 12

ID W27484 standard; Protein; 377 AA.
 AC W27484.
 DT 14-APR-1998 (first entry)
 DE Human MCP.
 KW Human; MCP; surface; homologous complement restriction factor;
 KW HCRF; medical apparatus; medical dressing; surgical equipment;
 KW diagnostic kit; prevention; purification device; reduction;
 KW complement activation.
 PS Homo sapiens.
 PN W09735886-A1.
 PD 02-OCT-1997.
 PF 12-MAR-1997; G00684.
 PR 31-OCT-1996; GB-022694.
 PR 22-MAR-1996; GB-006073.
 PR 28-MAR-1996; GB-006516.
 PA (IMOT-) IMOTRAN LTD.
 PI Watkins NJ.
 DR WPI; 97-489571/45.
 DR N-PSDB; T90306.
 PT Surfaces which prevent or reduce complement activation - having
 PT surface bound homologous complement restriction factor
 PS Example 1; Fig 1A; 50pp; English.
 CC The present sequence is human MCP, which was used in the
 CC development of a novel surface bound to a homologous complement
 CC restriction factor (HCRF), where the surface is not a surface to
 CC which HCRF binds in vivo. The surface can be used in medical
 CC apparatus (e.g. extra-corporeal circulation systems, tubing,
 CC valves, membranes, pumps, oxygenators, catheters, cannulas, fluid
 CC reservoirs or prostheses), medical dressings, surgical equipment,
 CC diagnostic kits (e.g. kits for determining whether a patient has
 CC an abnormality which prevents the complement system from working
 CC normally or from being properly regulated) and purification
 CC devices (e.g. for purifying complement components). The surface
 CC can be used to reduce or prevent activation of complement, and to
 CC reduce morbidity due to complement activation.
 SQ Sequence 377 AA;

Query Match 16.5%; Score 335; DB 1; Length 377;
 Best Local Similarity 31.0%; Pred. No. 5.29e-24;

Matches 81; Conservative 54; Mismatches 99; Indels 27; Gaps 22;

Db 33 DACEE-PTPEFAMELLIK-KPYEIEGERVDYCKKGYFIPLATHHTIDRHHMLPVS 90
 QY 1 EDCNEIPRRNTEILGNSWSDQTYPEGTQAIKCRPGYSLGNV-I-MVC-RKGEWALN 57
 Db 91 D-DACYRETCPYIRD-PLNGQAVPANGTYEFGYOMHFCINEGYLLIGETILCELKGSVA 148
 QY 58 PLKCCQKRCRPGHGDTPFGFTLLGNGVFEYGVKAVYTNEGQYLLIGE-INYRE-CDTGG 115
 Db 149 IMSGRPICKEKYLCTPPPKIKNGKHTFSEYEV-FEY-L-DAVYSCDPAAGPDPFSLIGE 205
 QY 116 -WTNDIPICEVVKCLPVTAPENCKIYSSAMEPDRHYHFGQAVFYCN-S-G---YKIEGD 169

Db 206 STIYCGNSVWSRAAPCEKRVKCRFPVYENGKQISGFGKKFYKATVME--CDKGF-YL 262
 QY 170 EEMHCSDGDFWMSKEKRCVCEISCKSPDVINGSPIS---QKIYKENERFYQKCNMGYEYS 226
 Db 263 DGSDDIVCDNSNTWDPVPVKC 283
 QY 227 ERGDA-VC-TESGM-RPLPSC 244

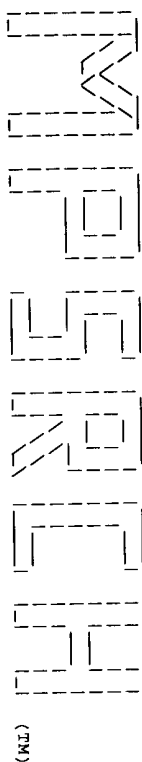
RESULT 13

ID R10924 standard; Protein; 384 AA.
 AC R10924;
 DT 09-MAY-1991 (first entry)
 DE Human membrane cofactor protein isoform.
 KW Mammalian membrane cofactor protein; complement activity;
 KW inflammation; autoimmune diseases; tissue injury.
 FH Key
 FT region
 FT 1..251
 FT /label= short consensus repeats (4)
 FT region
 FT 252-266
 FT /label= serine-threonine rich region, STB
 FT region
 FT 267..280
 FT /label= STC
 FT region
 FT 281..294
 FT /label= unknown region
 FT region
 FT 295..326
 FT /label= hydrophobic region
 FT region
 FT 327..350
 FT /label= CYN2 cytoplasmic tail region
 PN W09102002-A.
 PD 21-FEB-1991.
 PF 20-JUL-1990; U04107.
 PR 21-JUL-1989; US-384210.
 PR 19-APR-1990; US-510709.
 PA (UNITW) UNIV OF WASHINGTON.
 PI Atkinson JP;
 DR WPI; 91-073491/10.
 DR N-PSDB; Q10864.
 PT New recombinant mammalian membrane co-factor protein - for
 PT treating diseases with altered complement activity e.g.
 PT inflammatory and auto-immune conditions
 PS Disclosure: fig 1; 38pp; English.
 CC This human isoform of a membrane cofactor protein (MCP) is useful
 CC in a pharmaceutical compsn. for the treatment of inflammatory and
 CC autoimmune diseases mediated by excess or misdirection of comple-
 CC ment activity, e.g. rheumatoid arthritis or multiple sclerosis.
 CC Protection against tissue injury caused by e.g. myocardial in-
 CC agnation this MCP can be used in autoimmune disease diagnosis to
 CC predict the probability of recurrent miscarriages by testing for
 CC MCP levels in the placenta.
 CC See also Q10865-66 and R10927.
 SQ Sequence 384 AA;

Query Match 16.5%; Score 335; DB 1; Length 384;
 Best Local Similarity 31.0%; Pred. No. 5.29e-24;

Matches 81; Conservative 54; Mismatches 99; Indels 27; Gaps 22;

Db 33 DACEE-PTPEFAMELLIK-KPYEIEGERVDYCKKGYFIPLATHHTIDRHHMLPVS 90
 QY 1 EDCNEIPRRNTEILGNSWSDQTYPEGTQAIKCRPGYSLGNV-I-MVC-RKGEWALN 57
 Db 91 D-DACYRETCPYIRD-PLNGQAVPANGTYEFGYOMHFCINEGYLLIGETILCELKGSVA 148
 QY 58 PLKCCQKRCRPGHGDTPFGFTLLGNGVFEYGVKAVYTNEGQYLLIGE-INYRE-CDTGG 115
 Db 149 IMSGRPICKEKYLCTPPPKIKNGKHTFSEYEV-FEY-L-DAVYSCDPAAGPDPFSLIGE 205
 QY 116 -WTNDIPICEVVKCLPVTAPENCKIYSSAMEPDRHYHFGQAVFYCN-S-G---YKIEGD 169
 Db 206 STIYCGNSVWSRAAPCEKRVKCRFPVYENGKQISGFGKKFYKATVME--CDKGF-YL 262
 QY 170 EEMHCSDGDFWMSKEKRCVCEISCKSPDVINGSPIS---QKIYKENERFYQKCNMGYEYS 226



Release 3.1a John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

Msearch.p protein - protein database search, using Smith-Waterman algorithm
Run on: Thu Jun 8 21:41:21 2000; Maspar time 16.86 Seconds
741.518 Million cell updates/sec
Tabular output not generated.

Title: >US-09-316-163-10
Description: (1-265) from US09316163.pep
Perfect score: 2029
Sequence: 1 EDCNELPPRRNTEILTGSWS.....EKSCDNPYIPNGDYSPLRIK 265

Scoring table: PAM 150
Gap 11

Searched: 142080 seqs, 47172406 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: 1:pir1 2:pir2 3:pir3 4:pir4

Statistics: Mean 43.012; Variance 68.118; scale 0.631

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	2029	100.0	449	1 NBHUS	complement factor H p	0.00e+00
2	2029	100.0	1231	1 NBHUS	complement factor H p	0.00e+00
3	1439	70.9	1234	1 NBHUS	complement factor H p	0.00e+00
4	952	46.9	669	2 S65551	factor H - bovine (fr	9.39e-211
5	587	28.9	1053	2 S46199	probable complement r	1.11e-70
6	406	20.0	597	1 S53711	C4b alpha chain prec	2.94e-66
7	388	19.1	263	1 WNVZSP	apolipoprotein H homo	8.65e-65
8	382	18.3	597	1 NBHUC4	C4b-binding protein a	4.20e-62
9	371	18.3	345	1 NBMS	apolipoprotein H prec	4.20e-62
10	371	18.3	558	2 S57953	C4b protein alpha ch	1.06e-58
11	357	17.6	263	1 C36838	complement control pr	9.92e-58
12	353	17.4	560	2 T16833	hypothetical protein	1.48e-55
13	344	17.0	830	2 A30359	p-selectin precursor	7.88e-55
14	341	16.8	302	1 WMBE1E	secretory complement	7.88e-55
15	341	16.8	360	1 WMBE2E	membrane-bound comple	7.88e-55
16	341	16.8	469	1 NBMS4	C4b-binding protein a	7.88e-55
17	341	16.8	788	2 I53821	p-selectin - rat	4.16e-54
18	338	16.7	349	2 G02913	sperm CD46 - human (f	4.16e-54
19	338	16.7	369	2 I57998	membrane cofactor pro	4.16e-54
20	338	16.7	610	1 I46001	C4b-binding protein a	2.39e-54
21	339	16.7	768	2 A42755	p-selectin precursor	3.82e-53
22	334	16.5	345	1 NBBO	apolipoprotein H prec	3.82e-53
23	335	16.5	377	2 I54479	membrane cofactor pro	2.19e-53

RESULT ENTRY	1	ALIGNMENTS
TITLE	NBHUHS	#type complete
ALTERNATE_NAMES	complement factor H precursor, short splice form - human	
ORGANISM	complement factor H-related protein; complement protein H	
DATE	#formal name Homo sapiens #common name man	
ACCESSIONS	31-Dec-1993 #sequence revision 23-Feb-1996 #text change	
REFERENCE	22-Jun-1999	
#authors	S03013; B60238; A27877; A61103; A26505; S10479	
#journal	S00254	
#title	Ripoche, J.; Day, A.J.; Harris, T.J.R.; Sim, R.B.	
	Biochem. J. (1988) 249:593-602	
	The complete amino acid sequence of human complement factor H.	
#cross-references	MOID:88134059	
#accession	S03013	
#molecule-type	mRNA	
#residues	1-449 #label RIP	
#cross-references	EMBL:X07523; EMBL:Y00716; NID:g32492;	
#note	PIDN:CA30403.1; PID:g758073	
	part of this sequence, including the amino end of the	
	mature protein was confirmed by protein sequencing	
	402-Tyr was also found	
#note	not compared with conceptual translation	
#accession	B60238	
#molecule-type	mRNA	
#residues	1-33,434-449 #label EST	
#note	only portions of this 1.8 kilobase mRNA were sequenced	
REFERENCE	A27877	
#authors	Schulz, T.F.; Schwaible, W.; Stanley, K.K.; Weiss, E.	
#journal	Dierich, M.P.	
#title	Eur. J. Immunol. (1986) 16:1351-1355	
	Human complement factor H: isolation of cDNA clones and	
	partial cDNA sequence of the 38-kDa tryptic fragment	
	containing the binding site for C3b.	
#cross-references	MOID:87054207	
#accession	A27877	
#molecule-type	mRNA	
#residues	'U', '55-401', 'V', '403-449 #label SCH	
#note	an additional nucleotide present within the codon for	
	Glu-310 was thought to be a cloning artifact and was	
	ignored in translation	

```

REFERENCE
#authors Schwaebble, W.; Zwirner, J.; Schulz, T.F.; Linke, R.P.;
#journal Eur. J. Immunol. (1987) 17:1485-1489
#title Human complement factor H: expression of an additional
#cross-references GDB:88055295
#accession A61103
#status not compared with conceptual translation
#molecule_type mRNA
#residues 27-76 #label SC2
#note this is a partial sequence of an alternatively spliced
1.8 kilobase mRNA that is translated to yield a 43 k
form related to factor H

REFERENCE
#authors Sim, R.B.; Discipio, R.G.;
#journal Biochem. J. (1982) 205:285-293
#title Purification and structural studies on the complement-system
control protein beta-1-H (factor H).
#cross-references MIM:83048213
#accession A26505
#molecule_type protein
#residues 19-20,'Q',22-29,'V',31-33,'Q',35 #label SIM

REFERENCE
#authors Barlow, P.N.; Norman, D.G.; Steinkasserer, A.; Horne, T.J.;
#journal Biochemistry (1992) 31:3626-3634
#title Solution structure of the fifth repeat of factor H: A second
example of the complement control protein module.
#cross-references MIM:9232649
#accession S10479
#molecule_type protein
#authors Kristensen, T.; Metzel, R.A.; Tack, B.F.
#journal J. Immunol. (1986) 136:3407-3411
#title Structural analysis of human complement protein H: homology
with C4b binding protein, beta(2)-glycoprotein I, and the
Ba fragment of B.
#cross-references MIM:86169701
#accession S10479
#molecule_type mRNA
#residues 226-401,'Y',403-449 #label KRI
#cross-references GB:M12383; NID:9180472; PID:AAA52013.1; PID:9180473
COMMENT Factor H has also been found bound to cell membranes in an unknown
manner. However, it has at least one cell attachment site motif
in repeat 4.
COMMENT Alternative transcripts of 4.3, 1.8, and 1.4 kilobases are expressed
in liver. See also PIR:NBH0H.
GENETICS
#gene GDB:HP1; HP
#cross-references GDB:1120041; OMIM:134370
#map_position 1q32-1q32
GENETICS
#gene GDB:HP2; HP
#cross-references GDB:129095
#map_position 1q32-1q32
#note the correspondence between the two loci and the sequences
indicated is unclear; factor H has been reported to have
several allelic forms

FUNCTION
#description a cofactor in the inactivation of C3b by serine proteinase I;
also increases the rate of dissociation of the C3bBb
complex (C3 convertase) and the (C3b)bb complex (C5
convertase) in the alternative complement pathway
#pathway complement alternate pathway
#supernaturally complement factor H; complement factor H repeat
homology
#alternative splicing: complement alternate pathway;
glycoprotein; plasma

KEYWORDS
#domain signal sequence #status predicted #label SIG\
#product complement factor H, short splice form #status
experimental #label MAT\
#domain complement factor H repeat homology #label FH01\

FEATURE
1-18 #domain signal sequence #status predicted #label SIG\
19-449 #product complement factor H, short splice form #status
experimental #label MAT\
21-80 #domain complement factor H repeat homology #label FH01\

```

```

85-141 #domain complement factor H repeat homology #label FH02\
146-205 #domain complement factor H repeat homology #label FH03\
210-262 #domain complement factor H repeat homology #label FH04\
246-248 #region cell attachment (R-G-D) motif\
267-320 #domain complement factor H repeat homology #label FH05\
325-385 #domain complement factor H repeat homology #label FH06\
389-442 #domain complement factor H repeat homology #label FH07\
21-66,92-80,95-129,
114-141,146-192,
178-205,210-251,
237-262,267-309,
294-320,325-374,
357-385,389-431,
416-442
#disulfide_bonds #status predicted\
#binding_site carbohydrate (Asn) (covalent) #status
absent

SUMMARY
#length 449 #molecular_weight 51007 #checksum 6077
Query Match 100.0%; Score 2029; DB 1; Length 449;
Best local similarity 100.0%; Pred. No. 0.00e+00;
Matches 265; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 19 EDCNELPPRRNTIELTGSMSDQTYEGTQATYKCPGYSIGANYIMCRKGWALNPLR 78
1 EDCNELPPRRNTIELTGSMSDQTYEGTQATYKCPGYSIGANYIMCRKGWALNPLR 60
Db 79 KCKRPGCGHDPGPTGTLTGNGVFEYGVAVYTCNKGVLGELINRECDTGTNDI 138
61 KCKRPGCGHDPGPTGTLTGNGVFEYGVAVYTCNKGVLGELINRECDTGTNDI 120
Db 139 PICEVVKCLPYTAPENGRIVSAMPDREYHFGQAVFVCSGKIGDEMRHSDGFW 198
121 PICEVVKCLPYTAPENGRIVSAMPDREYHFGQAVFVCSGKIGDEMRHSDGFW 180
Db 199 SKERPKVEICCKSPDYVNGSPISOKIITYENEFQYKCNNGEYSESGDVAFTESGWRP 258
161 SKERPKVEICCKSPDYVNGSPISOKIITYENEFQYKCNNGEYSESGDVAFTESGWRP 240
Db 259 LPSCERKSCDNPYIPNGDYSPLRIK 283
241 LPSCERKSCDNPYIPNGDYSPLRIK 265
Qy 241 LPSCERKSCDNPYIPNGDYSPLRIK 265

RESULT 2
ENTRY NBH0H #type complete
TITLE complement factor H precursor, long splice form - human
ORGANISM #formal_name Homo sapiens #common_name man
DATE 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change
22-Jun-1999
ACCESSION S00254; A54726; A61565; A26505; I72654; S66298
#authors Ripoche, J.; Day, A.J.; Harris, T.J.R.; Sim, R.B.
#journal Biochem. J. (1988) 249:593-602
#title The complete amino acid sequence of human complement factor
H.
#cross-references MIM:88134059
#accession S00254
#molecule_type mRNA
#residues 1-1231 #label RIP
#cross-references EMBL:Y00716; NID:931964; PID:CAA68704.1; PID:931965
#note 402-Tyr was also found
#note parts of this sequence, including the amino and carboxyl
ends of the mature protein, were confirmed by protein
sequencing

REFERENCE
#authors Estaller, C.; Schwaebble, W.; Dierich, M.; Weiss, E.H.
#journal Eur. J. Immunol. (1991) 21:799-802
#title Human complement factor H: two factor H proteins are derived
from alternatively spliced transcripts.
#cross-references MIM:91184292
#accession A60238
#status not compared with conceptual translation
#molecule_type mRNA

```

##residues 1-56;1177-1231 ##label EST
 ##note only portions of this 4.3 kilobase mRNA were sequenced

REFERENCE
 #authors A54726
 Day, A.J.; Ripoché, J.; Lyons, A.; McIntosh, B.; Harris, T.J.R.; Sim, R.B.
 #journal Biosci. Rep. (1987) 7:201-207
 #title Sequence analysis of a cDNA clone encoding the C-terminal end of human complement factor H.

#cross-references M01D:8025472
 #accession A54726
 #status not compared with conceptual translation
 #molecule_type mRNA
 ##residues 'DPRN',579-1231 ##label DAY
 ##cross-references GB:M17517; NID:9180497; PIDN:AA52016.1; PID:9180498
 ##note parts of this sequence were determined by protein sequencing

REFERENCE
 #authors A61565
 Ripoché, J.; Day, A.J.; Willis, A.C.; Belt, K.T.; Campbell, R.D.; Sim, R.B.
 #journal Biosci. Rep. (1986) 6:65-72
 #title Partial characterization of human complement factor H by protein and cDNA sequencing: homology with other complement and non-complement proteins.

#cross-references M01D:86188123
 #accession A61565
 #status not compared with conceptual translation
 #molecule_type mRNA
 ##residues 'METERNHLNKKI',1050-1057,'T',1059-1102 ##label RI2

REFERENCE
 #authors A26505
 Sim, R.B.; Discipio, R.G.
 #journal Biochem. J. (1982) 205:285-293
 #title Purification and structural studies on the complement-system control protein beta-1-H (factor H).

#cross-references M01D:83048213
 #accession A26505
 #molecule_type protein
 ##residues 19-20,'Q',22-29,'V',31-33,'Q',35 ##label SIM

REFERENCE
 #authors A44551
 Barlow, P.N.; Norman, D.G.; Steinkasserer, A.; Horne, T.J.; Pearce, J.; Driscoll, P.C.; Sim, R.B.; Campbell, I.D.
 #journal Biochemistry (1992) 31:3626-3634
 #title Solution structure of the fifth repeat of factor H: A second example of the complement control protein module.

#cross-references M01D:9223649
 #contents annotation: NMR structure determination, residues 264-292
 A49224

REFERENCE
 #authors Norman, D.G.; Barlow, P.N.; Baron, M.; Day, A.J.; Sim, R.B.; Campbell, I.D.
 #journal J. Mol. Biol. (1991) 219:717-725
 #title Three-dimensional structure of a complement control protein module in solution.

#cross-references M01D:91278097
 #contents annotation: NMR structure determination, residues 927-985
 I56100

REFERENCE
 #authors Estélier, C.; Koistinen, V.; Schwaebler, W.; Dierich, M.P.; Weiss, E.H.
 #journal J. Immunol. (1991) 146:3190-3196
 #title Cloning of the 1.4-kb mRNA species of human complement factor H reveals a novel member of the short consensus repeat family related to the carboxy terminal of the classical 150-KDa molecule.

#cross-references M01D:91201892
 #accession I72654
 #status translated from GB/EMBL/DBJ
 #molecule_type mRNA

REFERENCE
 #residues 1047-1231 ##label RES
 ##cross-references GB:M65594; NID:9183766; PIDN:AA55948.1; PID:9183767
 S66298

REFERENCE
 #authors Carion, J.A.; Bates, R.C.; Smith, A.I.; Tetz, T.; Arellano, A.; Gordon, D.L.; Burns, G.F.
 #journal Biochim. Biophys. Acta (1996) 1289:305-311
 #title Factor H co-purifies with thrombospondin isolated from platelet secretate.

#cross-references M01D:96205365
 #accession S66298
 #status preliminary
 #molecule_type protein
 ##residues 411-419;574-578;580-582 ##label CAR
 #comment Factor H has also been found bound to cell membranes in an unknown manner. However, it has at least one cell attachment site motif in repeat 4.

#comment Alternative transcripts of 4.3, 1.8, and 1.4 kilobases are expressed in liver. See also PIR:NBH05.

GENETICS
 #gene GDB:HF1; HF
 #cross-references GDB:120041; OMIM:134370
 #map_position 1q32-1q32
 HF2
 GENETICS
 #gene GDB:HF2; HF
 #cross-references GDB:129095
 #map_position 1q32-1q32
 #note the correspondence between the two loci and the sequences indicated is unclear; factor H has been reported to have several allelic forms

FUNCTION
 #description a cofactor in the inactivation of C3b by serine proteinase I; also increases the rate of dissociation of the C3bb complex (C3 convertase) and the (C3b)bb complex (C5 convertase) in the alternative complement pathway
 #pathway complement alternate pathway
 #superfamily complement factor H; complement factor H repeat homology
 #alternative splicing; complement alternate pathway; glycoprotein; plasma

KEYWORDS
 #pathway
 CLASSIFICATION
 #keywords

FEATURE
 1-18
 19-1229
 19-449

21-80
 85-141
 146-205
 210-262
 246-248
 267-320
 325-385
 389-442
 448-505
 509-564
 569-623
 630-684
 691-744
 753-803
 811-864
 870-926
 931-984
 988-1043
 1048-1102
 1109-1163
 1167-1228
 21-66;52-80;85-129,
 114-141;146-192,
 178-205;210-251,
 237-267;267-309,
 294-320;325-374,
 357-385;389-431,
 416-442;448-494,
 477-505;509-553,
 536-564;569-611,
 597-623;630-673,
 659-684;691-733,
 719-744;753-792,
 781-803;811-853,
 839-864;870-915,
 901-926;931-973,

#domain complement factor H repeat homology #label FH01
 #domain complement factor H repeat homology #label FH02
 #domain complement factor H repeat homology #label FH03
 #domain complement factor H repeat homology #label FH04
 #region cell attachment (R-G-D) motif
 #domain complement factor H repeat homology #label FH05
 #domain complement factor H repeat homology #label FH06
 #domain complement factor H repeat homology #label FH07
 #domain complement factor H repeat homology #label FH08
 #domain complement factor H repeat homology #label FH09
 #domain complement factor H repeat homology #label FH10
 #domain complement factor H repeat homology #label FH11
 #domain complement factor H repeat homology #label FH12
 #domain complement factor H repeat homology #label FH13
 #domain complement factor H repeat homology #label FH14
 #domain complement factor H repeat homology #label FH15
 #domain complement factor H repeat homology #label FH16
 #domain complement factor H repeat homology #label FH17
 #domain complement factor H repeat homology #label FH18
 #domain complement factor H repeat homology #label FH19
 #domain complement factor H repeat homology #label FH20

Query Match	Best Local Similarity	Score	DB 1	Length	Length	#status
Matches	265	Conservative	0	Mismatches	0	Gaps
19	EDCNELPPRRNTEIITGSMDSOTVEEGQALYKCRPGKRSIGNTVMCRGGEWALNPLR	78	100.0%	2029	DB 1	Length 1231
1	EDCNELPPRRNTEIITGSMDSOTVEEGQALYKCRPGKRSIGNTVMCRGGEWALNPLR	60	100.0%	2029	DB 1	Length 1231
79	KCKRRCGHPGDPPTGFTLLTGAVNFEVGAAYTCNEDGYOLLGEINRYRCDDTDGWTNDI	138	100.0%	2029	DB 1	Length 1231
61	KCKRRCGHPGDPPTGFTLLTGAVNFEVGAAYTCNEDGYOLLGEINRYRCDDTDGWTNDI	120	100.0%	2029	DB 1	Length 1231
139	PICEVVKCLPYAPAPNGKIVSSAMEPDRHFYFGQAVFVPCNSGKILGDDEMHCSDDGFV	198	100.0%	2029	DB 1	Length 1231
121	PICEVVKCLPYAPAPNGKIVSSAMEPDRHFYFGQAVFVPCNSGKILGDDEMHCSDDGFV	180	100.0%	2029	DB 1	Length 1231
199	SKEKPKVEISCKSPDVINGSPISQKIIYKENERFQYKCMNGEYSPRGDAVTESGWRP	258	100.0%	2029	DB 1	Length 1231
181	SKEKPKVEISCKSPDVINGSPISQKIIYKENERFQYKCMNGEYSPRGDAVTESGWRP	240	100.0%	2029	DB 1	Length 1231
259	LPCSEKSCDNPYIPNGYSPLRK	283	100.0%	2029	DB 1	Length 1231
241	LPCSEKSCDNPYIPNGYSPLRK	265	100.0%	2029	DB 1	Length 1231
3	NBMSH	#type complete				
ENTRY	complement factor H precursor - mouse					
TITLE	protein beta-1-H					
ALTERNATE_NAMES	#formal_name Mus musculus #common_name house mouse					
ORGANISM	30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change					
DATE	22-Jun-1999					
ACCESSIONS	A26154; I49711; I49728					
REFERENCE	A26154					
#authors	Kristensen, T.; Tack, B.F.					
#journal	Proc. Natl. Acad. Sci. U.S.A. (1986) 83:3963-3967					
#title	Murine protein H is comprised of 20 repeating units, 61 amino acids in length.					
#cross-references	MUTID:86233353					
#accession	A26154					
#molecule_type	mRNA					
##residues	1-1234	##label	KRI			
##cross-references	GB:M12650; NID:g193724; PIDN:AAA37759.1; PID:g387181					
REFERENCE	I49711					
#authors	Natsuumi, Sakai, S.; Nonaka, M.; Nonaka, M.; Harada, Y.					
#journal	J. Immunol. (1990) 144:358-362					
#title	Demonstration of an unusual allelic variation of mouse factor H by the complete cDNA sequence of the H.2 allotype.					
#cross-references	MUTID:90111033					
#accession	I49711					
##status	preliminary: translated from GB/EMBL/DBJ					
##molecule_type	mRNA					
##residues	1-18	##label	RES			
##cross-references	GB:M1979; NID:g193726; PIDN:AAA37762.1; PID:g193729					
REFERENCE	I49728					
#authors	Munoz-Cano, P.; Tack, B.F.; Vik, D.P.					

```

#journal      Biochemistry (1989) 28:9891-9897
#title        Analysis of complement factor H mRNA expression:
              Dexamethasone and IFN-gamma increase the level of H in L
              cells.
#cross-references M01D:90148935
#accession    149728
#status       Preliminary; translated from GB/EMBL/DBJ
#molecule_type mRNA
#residues     1-19 ##label RE2
#cross-references GB:J02891; NID:9193805; PIDN:AAA37795.1; PID:g553926
COMMENT       Two codominant alleles of factor H are present in mice.
              Factor H functions as a cofactor in the inactivation of C3b by
              serine protease I and also increases the rate of dissociation
              of the C3bb complex (C3 convertase) and the (C3b)nbb complex (C3
              convertase) in the alternative complement pathway.
GENETICS
#map_position 1
CLASSIFICATION #superfamily complement factor H; complement factor H repeat
               homology
KEYWORDS        complement alternate pathway; duplication; glycoprotein;
               plasma
FEATURE
1-18            #domain signal sequence #status predicted #label SIG\
19-1234         #product complement factor H #status predicted #label
               MP\
21-80           #domain complement factor H repeat homology #label FH01\
85-141          #domain complement factor H repeat homology #label FH02\
145-205         #domain complement factor H repeat homology #label FH03\
210-262         #domain complement factor H repeat homology #label FH04\
245-248         #region cell attachment (R-G-D) motif\
267-3320        #domain complement factor H repeat homology #label FH05\
325-385         #domain complement factor H repeat homology #label FH06\
389-442         #domain complement factor H repeat homology #label FH07\
445-505         #domain complement factor H repeat homology #label FH08\
509-554         #domain complement factor H repeat homology #label FH09\
569-622         #domain complement factor H repeat homology #label FH10\
629-663         #domain complement factor H repeat homology #label FH11\
699-743         #domain complement factor H repeat homology #label FH12\
755-802         #domain complement factor H repeat homology #label FH13\
809-861         #domain complement factor H repeat homology #label FH14\
867-931         #domain complement factor H repeat homology #label FH15\
935-999         #domain complement factor H repeat homology #label FH16\
994-1048        #domain complement factor H repeat homology #label FH17\
1053-1107       #domain complement factor H repeat homology #label FH18\
1114-1136       #domain complement factor H repeat homology #label FH19\
1172-1133       #domain complement factor H repeat homology #label FH20\
21-66,52-80,85-129,
114-141,146-192,
178-205,210-251,
237-262,267-309,
294-320,325-374,
357-385,389-431,
416-442,448-494,
477-505,509-553,
536-564,569-610,
597-622,629-672,
658-683,690-732,
718-743,752-791,
780-802,808-850,
836-861,867-920,
906-931,936-978,
964-989,994-1037,
1023-1048,
1053-1096,
1082-1107,
1114-1157,
1143-1168,
1172-1233,
1206-1233
676,721,773,801,
1030,1061,1225
#disulfide_bonds #status predicted\
#binding_site carbohydrate (asn) (covalent) #status
              predicted
#length 1234 #molecular-weight 139081 #checksum 3676
SUMMARY

```


Query Match 70.9%; Score 1439; DB 1; Length 1234;
 Best Local Similarity 67.0%; Pred. No. 0.00e+00;
 Matches 177; Conservative 39; Mismatches 48; Indels 0; Gaps 0;

Db 19 EDCGPPRESEILSGMSSEOLYEGTQATYKCPGRRTGTITVYCKNGKMWASNP8R 78
 1 EDCNLPRLRRNTLITGMSDQTYETGTOAIYKCPGRSLGNVIMCRKGEWALNPLR 60
 Db 79 ICRKPCGHPGDTPEGSERLAVGSOFEFGAVYTCDDGTYOLLGEIDYREGADGWINDI 138
 61 KQKRPCHPGDTPGTFTLTGNAVFEYGAAYTCNNGYQLLGEINRECDTGWINDI 120
 Db 139 PCEVYKCLPTLENGRIVSGAETDQYGVGVFECSNGEFTIEGKIEHSCENGIM 198
 121 PCEVYKCLPTVAPANGKIVSSAMEPDRYHFGQAVFVCSNGKIEGDEHHCSDGFW 180
 Db 199 SNEKPCVETLCTPRVENDGINVKPVYKENERYHCKKHGYPKRGDAVCTGSGMS 258
 181 SKEKPCVETLCTPRVENDGINVKPVYKENERYHCKKHGYPKRGDAVCTGSGMS 240
 Db 259 QPCEERKCSPIYINGITTPHRI 282
 241 LPSCERKSCDNPYIPNGDYSPLRI 264

RESULT 4
 ENTRY 565551 #type fragment
 TITLE factor H - bovine (fragment)
 ORGANISM #formal_name Bos primigenius taurus #common_name cattle
 DATE 10-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 13-Aug-1999

ACCESSIONS
 REFERENCE S65551
 #authors Soemes, C.J.; Day, A.J.; Sim, R.B.
 #journal Biochem. J. (1996) 315:323-331
 #title Prediction from sequence comparisons of residues of factor H involved in the interaction with complement component C3b.
 #cross-references MIMD:96202005
 #accession S65551
 #status preliminary; not compared with conceptual translation
 #molecule_type mRNA
 #residues 1-669 #label SOA
 CLASSIFICATION #superfamily complement factor H; complement factor H repeat homology

FEATURE
 55-114 #domain complement factor H repeat homology #label FHR1
 296-349 #domain complement factor H repeat homology #label FHR2
 355-412 #domain complement factor H repeat homology #label FHR3
 416-471 #domain complement factor H repeat homology #label FHR4
 476-530 #domain complement factor H repeat homology #label FHR5
 538-592 #domain complement factor H repeat homology #label FHR6
 599-651 #domain complement factor H repeat homology #label FHR6

SUMMARY
 #length 669 #checksum 8335

Query Match 46.9%; Score 952; DB 2; Length 669;
 Best Local Similarity 64.3%; Pred. No. 9.39e-211;
 Matches 117; Conservative 29; Mismatches 36; Indels 0; Gaps 0;

Db 7 LAEGNPFYKAVYTCDEGQYMGANFRCDTNGTNDIPCEVYKCLPTPEPENGKI 66
 80 LTGNAVFEYGAAYTCNNGYQLLGEINRECDTGWINDIPEVYKCLPTAPENGKI 139
 Db 67 FSDALPEQETTYGVGVFECSNGYMLDGPQIHCASAGVSAETPCVETLCTPRVILN 126
 140 VSSAMEPDRYHFGQAVFVCSNGKIEGDEHHCSDGFWFKSKPKCVETLCTPRVILN 199
 Db 127 GOAVLPKATYQNERVQRCAGFEYGGGRTVCTKSWTAPICIELTQCPRIIPGVY 186
 200 GSPISQKIIYKENERFOYCKNMGYEYSEKGDVCTESGWRPLPSCERKSCDNPYIPNGDY 259
 Db 167 RP 188

QY 260 SP 261

RESULT 5
 ENTRY S46199 #type complete
 TITLE probable complement regulatory plasma protein SBI - barred sand bass
 ORGANISM #formal_name Paralichthys nebulifer #common_name barred sand bass
 DATE 19-Mar-1997 #sequence_revision 06-Jun-1997 #text_change 13-Aug-1999

ACCESSIONS
 REFERENCE S46199
 #authors Dahmen, A.; Kaidoh, T.; Zupfel, P.F.; Gigli, I.
 #journal Biochem. J. (1994) 301:391-397
 #title Cloning and characterization of a cDNA representing a putative complement-regulatory plasma protein from barred sand bass (Paralichthys nebulifer).
 #cross-references MIMD:94318039
 #accession S46199
 #molecule_type mRNA
 #residues 1-1053 #label DAH1
 #experimental_source liver
 #accession S77894
 #cross-references EMBL:L21703; NID:9639894; PID:9639895
 #residues 526-532, 'X', 534-537, 809-817, 'X', 819-826 #label DAH2

GENETICS
 SBI
 CLASSIFICATION #superfamily complement factor H repeat homology
 KEYWORDS glycoprotein
 FEATURE
 89-145 #domain complement factor H repeat homology #label FH01
 334-389 #domain complement factor H repeat homology #label FH02
 450-502 #domain complement factor H repeat homology #label FH03
 569-624 #domain complement factor H repeat homology #label FH04
 682-738 #domain complement factor H repeat homology #label FH05
 743-802 #domain complement factor H repeat homology #label FH05
 935-989 #domain complement factor H repeat homology #label FH05
 993-1052 #domain complement factor H repeat homology #label FH06

SUMMARY
 #length 1053 #molecular_weight 11737 #checksum 8482

Query Match 28.9%; Score 587; DB 2; Length 1053;
 Best Local Similarity 37.1%; Pred. No. 3.83e-116;
 Matches 91; Conservative 41; Mismatches 101; Indels 12; Gaps 9;

Db 47 EASYDGRQVRYGCVNGY-S-GFEKLYCBESKWTGCA--KQDRSCGHPDQAFADFL 102
 21 DQTYEGTQAIYKCPGRYSLGNVIMCRKGEWALNPLRCKQKPCGHPDTPFGFTL 80
 Db 103 AEGNDFVGSKYVTCQKQYGVSRINRCAVAGMDGVVVCESQOC-PLIHVDNNVQY 161
 81 TCGNAVFEYGAAYTCNNGYQLLGEINRECDTGWINDIPEVYKCLPTAPENGKIY 140
 Db 162 ICG--PE-EATGNAVFRFSCSKSRSEILDSPELYCDERGDWSGVPCAKATCAIPIEN 218
 141 SSAMEPDRYHFGQAVFVCSNGYKI-EGDEHHCSDGFWFKSKPKCVETLCTPRVILN 199
 QY 219 GNVPGAIRKYKNDVLYHCECPRAFKHIDPSCICIQGKAEMSPPLCESIKCRITIDG 278
 200 GSPISQKIIYKENERFOYCKNMGYEYSEKGDV--CTESGWRPLPSCERKSCDNPYIPN 256
 Db 279 TRYEP 283
 QY 257 GDYSP 261

RESULT 6
 ENTRY S53711 #type complete
 TITLE CABP alpha chain precursor - rabbit
 ORGANISM #formal_name Oryctolagus cuniculus #common_name domestic rabbit
 DATE 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

REFERENCE	FEATURE	KEYWORDS	CLASSIFICATION	REFERENCE	FEATURE	KEYWORDS	CLASSIFICATION
#journal #title #accession ##molecule-type DNA ##residues ##cross-references GB:M35027; NID:g35317; PIDN:AAA47997.1; PID:g335344 #experimental_source strain Copenhagen A42531	1-263 ##label GOE Winslow, J.P.; Paoletti, E. Virology (1990) 179:517-563 Appendix to "The complete DNA sequence of vaccinia virus", B42504	Winslow, J.P.; Paoletti, E. Virology (1990) 179:517-563 Appendix to "The complete DNA sequence of vaccinia virus", B42504		#journal #title #cross-references MUID:91021027 #contents #note	Goebel, S.J.; Johnson, G.P.; Perkus, M.E.; Davis, S.W.; Winslow, J.P.; Paoletti, E. Virology (1990) 179:247-266 The complete DNA sequence of vaccinia virus. MUID:91021027 annotation; possible protein-coding frames neither amino acid nor nucleotide sequence is given #superfamily herpesvirus complement control protein; duplication; extracellular protein		
#domain signal sequence #status predicted #label SIG\	#product C4b-binding protein homolog #status predicted	#domain complement factor H repeat homolog #label FH1\		#domain complement factor H repeat homolog #label FH2\			
#domain complement factor H repeat homolog #label FH3\		#domain complement factor H repeat homolog #label FH4\					
#domain complement factor H repeat homolog #label FH5\							
#domain complement factor H repeat homolog #label FH6\							
#domain complement factor H repeat homolog #label FH7\							
#domain complement factor H repeat homolog #label FH8\							
#domain complement factor H repeat homolog #label FH9\							
#domain complement factor H repeat homolog #label FH10\							
#domain complement factor H repeat homolog #label FH11\							
#domain complement factor H repeat homolog #label FH12\							
#domain complement factor H repeat homolog #label FH13\							
#domain complement factor H repeat homolog #label FH14\							
#domain complement factor H repeat homolog #label FH15\							
#domain complement factor H repeat homolog #label FH16\							
#domain complement factor H repeat homolog #label FH17\							
#domain complement factor H repeat homolog #label FH18\							
#domain complement factor H repeat homolog #label FH19\							
#domain complement factor H repeat homolog #label FH20\							
#domain complement factor H repeat homolog #label FH21\							
#domain complement factor H repeat homolog #label FH22\							
#domain complement factor H repeat homolog #label FH23\							
#domain complement factor H repeat homolog #label FH24\							
#domain complement factor H repeat homolog #label FH25\							
#domain complement factor H repeat homolog #label FH26\							
#domain complement factor H repeat homolog #label FH27\							
#domain complement factor H repeat homolog #label FH28\							
#domain complement factor H repeat homolog #label FH29\							
#domain complement factor H repeat homolog #label FH30\							
#domain complement factor H repeat homolog #label FH31\							
#domain complement factor H repeat homolog #label FH32\							
#domain complement factor H repeat homolog #label FH33\							
#domain complement factor H repeat homolog #label FH34\							
#domain complement factor H repeat homolog #label FH35\							
#domain complement factor H repeat homolog #label FH36\							
#domain complement factor H repeat homolog #label FH37\							
#domain complement factor H repeat homolog #label FH38\							
#domain complement factor H repeat homolog #label FH39\							
#domain complement factor H repeat homolog #label FH40\							
#domain complement factor H repeat homolog #label FH41\							
#domain complement factor H repeat homolog #label FH42\							
#domain complement factor H repeat homolog #label FH43\							
#domain complement factor H repeat homolog #label FH44\							
#domain complement factor H repeat homolog #label FH45\							
#domain complement factor H repeat homolog #label FH46\							
#domain complement factor H repeat homolog #label FH47\							
#domain complement factor H repeat homolog #label FH48\							
#domain complement factor H repeat homolog #label FH49\							
#domain complement factor H repeat homolog #label FH50\							
#domain complement factor H repeat homolog #label FH51\							
#domain complement factor H repeat homolog #label FH52\							
#domain complement factor H repeat homolog #label FH53\							
#domain complement factor H repeat homolog #label FH54\							
#domain complement factor H repeat homolog #label FH55\							
#domain complement factor H repeat homolog #label FH56\							
#domain complement factor H repeat homolog #label FH57\							
#domain complement factor H repeat homolog #label FH58\							
#domain complement factor H repeat homolog #label FH59\							
#domain complement factor H repeat homolog #label FH60\							
#domain complement factor H repeat homolog #label FH61\							
#domain complement factor H repeat homolog #label FH62\							

```

REFERENCE
#residues 495-505, 'X', '507-510, 'X', '512-515' ##label S02
#note this peptide appears to bind protein S
#authors Dahlback, B.; Smith, C.A.; Muller-Eberhard, H.J.
#journal Proc. Natl. Acad. Sci. U.S.A. (1983) 80:13461-13465
#title Visualization of human C4b-binding protein and its complexes with vitamin K-dependent protein S and complement protein C4b.
#cross-references MUID:83221615
#contents annotation: electron microscopy: three-dimensional structure; ligand binding
#authors ASO, T.; Okamura, S.; Matsuguchi, T.; Sakamoto, N.; Sata, T.; Niho, Y.
#journal Biochem. Biophys. Res. Commun. (1991) 174:222-227
#title Genomic organization of the alpha chain of the human C4b-binding protein gene.
#cross-references MUID:91113199
#accession I52244
#status translation not shown; translated from GB/EMBL/DBJ
##molecule-type DNA
##residues 1-597 ##label ASO
#cross-references GB:M62486; NID:q190498; PIDN:AAA36506.1; PID:q190500
#comment C4b controls the classical pathway of complement activation. It binds as a cofactor to C3b/C4b inactivator (C3bINA), which then hydrolyzes the complement fragment C4b. It also accelerates the degradation of the C4bC2a complex (C3 convertase) by dissociating the complement fragment C2a.
#comment C4b occurs in plasma in two forms, both of which bind complement fragment C4b. The prevalent higher molecular weight form contains 7 alpha chains and one beta chain, which are linked by disulfide bonds. The beta chain binds the vitamin K-dependent plasma protein S. A minor form lacks the beta chain. Bound protein S is inactive as a cofactor for protein C inactivation of coagulation factors V and VIII.
#comment The molecule has a central body supporting seven tentacles (alpha chains), each with the binding site for C4b at the peripheral end.

GENETICS
#gene GDB:C4BPA
#cross-references GDB:120568; OMIM:120830
#map_position Iq3f-Iq32
#introns 48/1; 110/1; 143/2; 172/1; 236/1; 297/1; 362/1; 425/1; 482/1; 540/3
#comment octamer of seven alpha chains and one beta chain
#classification #superfamily C4b-binding protein alpha chain; complement factor H repeat homology
#keywords acute phase; chylomicron; complement pathway; duplication; glycoprotein; plasma

FEATURE
#domain signal sequence #status predicted #label SIG\
1-48
#product C4b-binding protein alpha chain #status predicted #label MAT\
49-597
#domain complement factor H repeat homology #label FH1\
50-108
#domain complement factor H repeat homology #label FH2\
113-170
#domain complement factor H repeat homology #label FH3\
175-334
#domain complement factor H repeat homology #label FH4\
239-294
#domain complement factor H repeat homology #label FH5\
299-360
#domain complement factor H repeat homology #label FH6\
364-422
#region complement C4b binding #status predicted\
381-404
#domain complement factor H repeat homology #label FH7\
426-480
#domain complement factor H repeat homology #label FH8\
484-538
#binding_site carboxyrate (Asn) (covalent) #status experimental\
221,506,528

SUMMARY
#length 597 #molecular-weight 67033 #checksum 6374

Query Match 18.8%; Score 382; DB 1; Length 597;
Best Local Similarity 31.5%; Pred. No. 8,65e-65;
Matches 79; Conservative 52; Mismatches 98; Indels 22; Gaps 20;

Db 67 TETRTKGTTLTKYLCIPYVRSHSTOTLNCSDGW-VYNTFCIT-KR-CRHPGELRNGQ 123
::: || | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
20 SPQYPRPGQAIKRCRPGY-RSLGNVIMWCK-GEWVNLPLKRCQKRPCHGPDPTFGT 77

```

Db 124 VEIKT-DL-SFGSOIEFSCSEGFLLIGSTSRCEVDGRGVGMSHPRLPOCEIYKCKPPDI 181
 QY 78 FTLLGNGVFEYGVAAVWCMNGEYOLLGELINTR-EC-DTD-GWTDIPICEVYKCLPVPAP 134
 Db 182 RNRG-HSG-E-ENFYAVGFSVTYSCDPRFSLIGHASISCTVENETIGVMPSPPTCKE 237
 QY 135 ENGRIVSSAMEPDREYHGAQVRFVNCNGYKIEGDEMHCS-D-D-GFWSKEKPKVEI 190
 Db 238 TCRKPDVSHGMSVGFQYVYKDTIYKCKGKGYLRSSVYIHCDDASKNPPSPACEPN 297
 QY 191 SCKSPDIVNSPISQ-KITIKENERFOYKCMWGYEYSEKDAVC-TESGMRP-LPSCEEK 247
 Db 298 SCINLPDIPHA 308
 QY 248 SCDN-PYIPNG 257

ENTRY 9
 TITLE NEMS
 ALTERNATE_NAMES #type complete
 50K serum glycoprotein; activated protein C-binding protein;
 beta-2-glycoprotein I
 ORGANISM #formal_name Mus musculus #common_name house mouse
 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change
 28-May-1999
 ACCESSIONS A43286; JC2243
 REFERENCE A43286
 #authors Nonaka, M.; Matsuda, Y.; Shirosaki, T.; Moriwaki, K.; Nonaka, M.; Natsunome-Sakai, S.
 #journal Genomics (1992) 13:1082-1087
 #title Molecular cloning of mouse beta-2-glycoprotein I and mapping of the gene to chromosome 11.
 #cross-references MIMD:92372000
 #accession A43286
 #molecule_type mRNA
 #residues 1-252, 'A', 254-277, 'N', 279-345 ##label NON
 #cross-references GB:D10056
 #note The authors translated the codon ACT for residue 253 as Ala and ATG for residue 278 as Asn
 sequence extracted from NCBI backbone (NCBIN:111791,
 NCBI:P:111794) and corrected to correspond with the
 nucleotide translation

REFERENCE JC2243
 #authors Sellar, G.C.; Steel, D.M.; Zafiroopoulos, A.; Seery, L.T.; Whitehead, A.S.
 #journal Biochem. Biophys. Res. Commun. (1994) 200:1521-1528
 #title Characterization, expression and evolution of mouse beta2-glycoprotein I (apolipoprotein H).
 #cross-references MIMD:94242017
 #accession JC2243
 #molecule_type mRNA
 #residues 1-251, 'R', 253-345 ##label SEL
 #cross-references GB:S70439; NID:g546780; PID:AAB30789.1; PID:g546781
 #experimental_source liver

GENETICS B2gpl
 #gene map_position 11
 CLASSIFICATION #superfamily apolipoprotein H; complement factor H repeat
 #homology
 #chylomicron; duplication; glycoprotein; HDL; heparin binding; lipid binding; monomer; plasma; VLDL

KEYWORDS
 FEATURE 1-19 #domain signal sequence #status predicted #label SIG\
 20-345 #product apolipoprotein H #status predicted #label MAT\
 23-79 #domain complement factor H repeat homology #label FH1\
 84-137 #domain complement factor H repeat homology #label FH2\
 142-200 #domain complement factor H repeat homology #label FH3\
 205-260 #domain complement factor H repeat homology #label FH4\
 264-325 #domain complement factor H repeat homology #label FH5\
 23-66, 51-79, 84-124, #domain complement factor H repeat homology #label FH6\
 110-137, 142-188, #domain complement factor H repeat homology #label FH7\
 174-200, 205-248, #domain complement factor H repeat homology #label FH8

234-260, 264-315, #disulfide_bonds #status predicted\
 300-307, 325-345, #binding site carbohydrate (Asn) (covalent) #status
 105, 117, 162, 183, predicted
 193

SUMMARY #length 345 #molecular_weight 38618 #checksum 2016

Query Match 18.3%; Score 371; DB 1; Length 345;
 Best Local Similarity 29.4%; Pred. No. 4, 20e-62;
 Matches 74; Conservative 49; Mismatches 110; Indels 19; Gaps 15;

Db 40 SYDPEQIVYSCRPYSGRGMRFRCPLTGM-PITLRL-CYPRVCPFGILENGIVRY 97
 QY 23 TPEGTQAIYKCRPGYSLGNVTK-VCR-KGEWALNPLKRCQKRCRPGHGDPPFGFTL 80
 Db 98 TS--FEYPRNIFACNPGF-FLNGTSSSKCTEGKMSPDIPACARTICPPVPKFLAL 153
 QY 81 TGGVFEYGVKAVYTCNEGYYOLLGEIYRCDITDG-WTNDIPICEVYKCLPVPAPENKRI 139
 Db 154 KDIRPSAGNSLXODIVYFCLPHFAMIGNDIVYCTEQGMWRL-PECLEVKCPFPPE 212
 QY 140 VSSAMEPDREYHGAQVRFVNCNGYKIEGDEMHCSDDGFWSKKPKVEISCK-SPDVI 198
 Db 213 NGYVNPAPKPVLLIKDKATPGF--CHETRYKLDGPEAECTRTGWSPLPCRE-SCKLPAVK 269
 QY 199 NG--S-PISQKITIKENEROYKCMNGEYSEKDAVCTESG-WRPLPSCSEKSCDNPYI 254
 Db 270 KATVLYIQMRVK 281
 QY 255 PNGD-YSPULRIK 265

RESULT 10
 ENTRY S57953 #type complete
 TITLE C4b protein alpha chain precursor - rat
 ORGANISM #formal_name Rattus norvegicus #common_name Norway rat
 DATE 13-Jan-1996 #sequence_revision 01-Mar-1996 #text_change
 20-Aug-1999

ACCESSIONS S57953
 REFERENCE S57953
 #authors Hillarp, A.; Thern, A.; Dahlback, B.
 #submission Submitted to the EMBL Data Library, July 1995
 #description Molecular cloning of rat C4b-binding protein alpha- and beta-chains: structural and functional relationships among human, bovine, rabbit, mouse and rat proteins.

#accession S57953
 #status preliminary
 #molecule_type mRNA
 #residues 1-558 ##label HITL
 #cross-references EMBL:250051; NID:g899379; PID:CAA90391.1; PID:g899380

CLASSIFICATION #superfamily C4b-binding protein alpha chain; complement factor H repeat homology

FEATURE 15-72 #domain complement factor H repeat homology #label FH1\
 77-134 #domain complement factor H repeat homology #label FH2\
 139-199 #domain complement factor H repeat homology #label FH3\
 204-258 #domain complement factor H repeat homology #label FH4\
 263-324 #domain complement factor H repeat homology #label FH5\
 338-386 #domain complement factor H repeat homology #label FH6\
 390-443 #domain complement factor H repeat homology #label FH7\
 447-501 #domain complement factor H repeat homology #label FH8

SUMMARY #length 558 #molecular_weight 62265 #checksum 9928

Query Match 18.3%; Score 371; DB 2; Length 558;
 Best Local Similarity 30.5%; Pred. No. 4, 20e-62;
 Matches 73; Conservative 57; Mismatches 87; Indels 22; Gaps 18;

Db 43 YNCRPGYSRASSQSLXCKPLGKW-QIN-IA-CVKSQRNPGDLONGKVEYKT-D-FLFG 97
 QY 32 YKCRPGY-RSLGNVIMCRK-GEWALNPLKRCQKRCRPGHGDIPFGFTLTLGNGVEYIG 89
 Db 98 SGIERSCEGTYILLGSSISYCEIOGKGVSWSDPLPECVIARCKGMPDIS-NGK--HNGRE 154

```

OY 90 KAVATCBGQOLLE-INTRECDTGD--WTNDIPICEVVK-LPVTAPENKIVSSAME 145
Db 155 EEF-FTYRSTVYKCDPDPFTLLGNASITCTVYVKNKTVGVSPSPPCERICMPVYLHGT 213
OY 146 PDREYHFGQAVRFVNCNGKIKGDEEMHS--D--GFMSKCKPACVEISCSPDVIINS 201
Db 214 INSGKHKYKKKDSRVFCQKGFVLRGSGVHICGADGSMSPVPCVELNSCTDIPDIPNA 272
OY 202 PISQ-KIITYKENERFYKCMNGEYSEKGDVAVC-TESGWRPLPSCSEKSC-DNPIYIPNG 257

RESULT 11
ENTRY 11
TITLE C36838 #type complete
#domain complement factor H repeat homology - variola virus (strain India-1967)
ALTERNATE_NAMES D12L protein
ORGANISM #formal_name variola virus
DATE 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
ACCESSIONS C36838
REFERENCE A36859
#authors Blinov, V.M.
#submission submitted to Genbank, November 1992
#description not shown.
#accession C36838
#molecule_type DNA
#residues 1-263 ##label BLI
#cross-references GB:X69198; NID:9456758; PIDN:CAA48953.1; PID:9297195
CLASSIFICATION #superfamily herpesvirus complement control protein; complement factor H repeat homology

FEATURE
21-81 #domain complement factor H repeat homology #label FH1\
86-143 #domain complement factor H repeat homology #label FH2\
148-201 #domain complement factor H repeat homology #label FH3\
206-261 #domain complement factor H repeat homology #label FH4\
SUMMARY #length 263 #molecular_weight 28789 #checksum 8771

Query Match 17.6%; Score 357; DB 1; Length 263;
Best Local Similarity 33.6%; Pred. No. 1,06e-58;
Matches 79; Conservative 38; Mismatches 95; Indels 23; Gaps 19;

Db 40 ANANYIGDTIEVLCLEPKRKMGPIYAKCTGTG--TL--FNOCIRKRCSPRDINGH 96
OY 20 SDOTYEGGTQAIYKCRGYS--LGNIVMCRKGEMALNPLRCQKRPCHPDTPFGT 77
Db 97 LDI-IGG-V-DEGSSITVSCSGYLLIGEKYSYCKLGSTGSKMNPAPICESVQDPPS 153
OY 78 FTLTGAVFEYGVKAVYTCNEGYYQLGEIN-Y-RECDTDG--WTNDIPICEVVK-LPVT 132
Db 154 IS-NGR--HNGYN-DF-YTDSVVTYSCNSGYSLIGNSGYLCSGE--WSPN-PTQCIYK 206
OY 133 APENKIVSSAMPEPDREYHFGQAVRFVNCNGKIKGDEEMHSCTDGPWSEKRPCEVISC 192
Db 207 PHFTIINGYISSGFKSYSTNDVDTCKYGYLLSGSSSTGSPGNTWQPELPKC 261
OY 193 KSPDIVNGSPIQ-KIITYKENERFYKCMNGEYSEKGDVAVC--WRP-LPSC 244

RESULT 12
ENTRY 12
TITLE T16833 #type complete
#domain complement factor H repeat homology - Caenorhabditis elegans
ORGANISM #formal_name Caenorhabditis elegans
DATE 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
ACCESSIONS T16833
REFERENCE T18586
#authors Gelsel, C.
#submission submitted to the EMBL Data Library, April 1996
#description The sequence of C. elegans cosmid T07H6.
#accession T16833
#status Preliminary; translated from GB/EMBL/DBJ
#molecule_type DNA

```

```

#residues 1-560 ##label GEI
#cross-references EMBL:U53344; NID:91255886; PID:91255889;
#experimental_source strain Bristol N2; clone T07H6
GENETICS
#gene CESP:T07H6.5
#map_position 10
#introns 14/1; 75/1; 102/3; 128/1; 186/1; 272/2; 326/1; 422/1;
#length 560 #molecular_weight 61619 #checksum 2418

SUMMARY
Query Match 17.4%; Score 353; DB 2; Length 560;
Best Local Similarity 32.2%; Pred. No. 9,92e-58;
Matches 68; Conservative 42; Mismatches 82; Indels 19; Gaps 15;

Db 117 AQWEPD-LR-CKAPACDPDGIENG--LRBDTFEYHHYKSCNPGFLVGS--RQ 170
OY 51 GEWALNPLRCKQKRPCHPDTPFGFTLLGNVFEYGVKAVYTCNEGYYQLGEINRE 110
Db 171 CSNGETNEPANCATKATCSRPSSPLHGKVVGSSL-T---YQ-S-VYIYSCDHGRLYGQ 224
OY 111 CDTDG-WTNDIPICEVYVCLVYTAPEKIVSSAMEPDREYHFGQAVRFVNCNGKIKG 169
Db 225 VQICLAEGINGNEPRCEETRCGYLPTLPNGYIEGSETSGAVALFRCLETMTHE-GA- 282
OY 170 EEMHCSDGGEFWSKRPCKVEISCKS-PDIVNGSPIQKIIITYKENERFYKCMNGEYSE 228
Db 283 SKAKMEDGWSAPIPRLA-SCRPHIONG 312
OY 229 GDAVCTESG-WR-PLPSCSEKSCDNPIYIPNG 257

RESULT 13
ENTRY 13
TITLE A30359 #type complete
#domain complement factor H repeat homology - human
ALTERNATE_NAMES P-selectin precursor - human
ORGANISM CD62 antigen; granule membrane protein 140
DATE 18-Oct-1989 #sequence_revision 30-Sep-1991 #text_change 29-Aug-1997
ACCESSIONS A30359
REFERENCE A30359
#authors Johnston, G.I.; Cook, R.G.; McEver, R.P.
#journal Cell (1989) 56:1033-1044
#title Cloning of GMP-140, a granule membrane protein of platelets and endothelium: sequence similarity to proteins involved in cell adhesion and inflammation.
#cross-references MIMD:89168432
#accession A30359
#molecule_type mRNA
#residues 1-830 ##label JOH
#cross-references GB:M25322
#note parts of this sequence, including the amino end of the mature protein, were confirmed by protein sequencing

GENETICS
#gene GDB:SELP: GRNP
#map_position 1422-1925
CLASSIFICATION #superfamily complement factor H repeat homology; EGF homology
KEYWORDS cell adhesion; glycoprotein; phosphohistidine; phosphoprotein; surface antigen; transmembrane protein

FEATURE
1-41 #domain signal sequence #status predicted #label SIG\
42-830 #product P-selectin #status experimental #label MAR\
163-194 #domain EGF homology #label EGF\
200-257 #domain complement factor H repeat homology #label FH01\
324-319 #domain complement factor H repeat homology #label FH02\
324-381 #domain complement factor H repeat homology #label FH03\
386-443 #domain complement factor H repeat homology #label FH04\
448-505 #domain complement factor H repeat homology #label FH05\
510-567 #domain complement factor H repeat homology #label FH06\
572-629 #domain complement factor H repeat homology #label FH07\
642-699 #domain complement factor H repeat homology #label FH08\

```


Fri Jun 9 10:53:38 2000

US-09-316-163-10.rpt

Page 11

QY 196 DVING-SPISOKIYKENERFOYKCNMGE-YSERGDVCTESGWRP-LPSC 244

Search completed: Thu Jun 8 21:41:41 2000
Job time : 20 secs.

RL Biochem. J. 205:285-293(1982).
 [6]
 RN STRUCTURE BY NMR OF 927-985 (SUSHI 16).
 RX MEDLINE: 91278097.
 RA Norman D.G., Barlow P.N., Baron M., Day A.J., Sim B., Campbell I.D.;
 RT "three-dimensional structure of a complement control protein module
 in solution."; J Mol. Biol. 219:717-725(1991).
 RN [7]
 RP STRUCTURE BY NMR OF 264-322 (SUSHI 5).
 RX MEDLINE: 92232649.
 RA Barlow P.N., Norman D.G., Steinkasserer A., Horne T.J., Pearce J.,
 RT "Discoil P.C., Sim B., Campbell I.D.;
 RT "Solution structure of the fifth repeat of factor H: a second example
 of the complement control protein module."; Biochemistry 31:3626-3634(1992).
 RL [8]
 RN STRUCTURE BY NMR OF 866-985 (SUSHI 15 AND 16).
 RX MEDLINE: 93323119.
 RA Barlow P.N., Steinkasserer A., Norman D.G., Kieffer B., Wiles A.P.,
 RT "Solution structure of a pair of complement modules by nuclear
 magnetic resonance."; J. Mol. Biol. 232:268-284(1993).
 CC -1- FUNCTION: FACTOR H FUNCTIONS AS A COFACTOR IN THE INACTIVATION OF
 C3B BY FACTOR I AND ALSO INCREASES THE RATE OF DISSOCIATION OF THE
 C3BB COMPLEX (C3 CONVERTASE) AND THE (C3)NBB COMPLEX (C5
 CONVERTASE) IN THE ALTERNATIVE COMPLEMENT PATHWAY.
 CC -1- SIMILARITY: CONTAINS 20 SUSHI (SCR) REPEATS.
 CC -1- CAUTION: REF.2 SEQUENCE DIFFERS FROM THAT SHOWN FROM POSITION
 341 ONWARD DUE TO A FRAMESHIFT.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: Y00716; CAA68704.1; -
 DR EMBL: X04697; CAB41739.1; ALT_FRAME.
 DR EMBL: M65294; AAA35948.1; -
 DR PIR: S00254; NBHCH.
 DR PIR: S00254; S00254.
 DR PIR: S03013; S03013.
 DR PDB: 1HCC; 15-APR-92.
 DR PDB: 1HFH; 15-JUL-93.
 DR PDB: 1HFI; 15-JUL-93.
 DR MIM: 134370; -
 DR PFM: PF00084; sush1.20.
 KW Complement alternate pathway; Plasma; Glycoprotein; Repeat; Sush1;
 KW Signal; 3d-structure; Polymorphism.
 FT SIGNAL 1 18
 FT CHAIN 19 1331
 FT DOMAIN 20 1229
 FT REPEAT 20 8
 FT REPEAT 84 142
 FT REPEAT 145 206
 FT REPEAT 209 263
 FT REPEAT 266 321
 FT REPEAT 324 386
 FT REPEAT 388 443
 FT REPEAT 447 506
 FT REPEAT 508 565
 FT REPEAT 568 624
 FT REPEAT 629 685
 FT REPEAT 690 745
 FT REPEAT 752 804
 FT REPEAT 810 865
 FT REPEAT 869 927
 FT REPEAT 930 985
 FT REPEAT 988 1044
 FT SUSHI 17.

FT REPEAT 1047 1103 SUSHI 18.
 FT REPEAT 1108 1164 SUSHI 19.
 FT REPEAT 1166 1229 SUSHI 20.
 FT DISULFID 21 66 BY SIMILARITY.
 FT DISULFID 52 80 BY SIMILARITY.
 FT DISULFID 85 129 BY SIMILARITY.
 FT DISULFID 114 141 BY SIMILARITY.
 FT DISULFID 146 192 BY SIMILARITY.
 FT DISULFID 178 205 BY SIMILARITY.
 FT DISULFID 210 251 BY SIMILARITY.
 FT DISULFID 237 262 BY SIMILARITY.
 FT DISULFID 267 309 BY SIMILARITY.
 FT DISULFID 294 320 BY SIMILARITY.
 FT DISULFID 325 374 BY SIMILARITY.
 FT DISULFID 357 385 BY SIMILARITY.
 FT DISULFID 389 431 BY SIMILARITY.
 FT DISULFID 416 442 BY SIMILARITY.
 FT DISULFID 448 494 BY SIMILARITY.
 FT DISULFID 477 505 BY SIMILARITY.
 FT DISULFID 509 553 BY SIMILARITY.
 FT DISULFID 536 564 BY SIMILARITY.
 FT DISULFID 569 611 BY SIMILARITY.
 FT DISULFID 597 623 BY SIMILARITY.
 FT DISULFID 630 673 BY SIMILARITY.
 FT DISULFID 659 684 BY SIMILARITY.
 FT DISULFID 691 733 BY SIMILARITY.
 FT DISULFID 719 744 BY SIMILARITY.
 FT DISULFID 753 792 BY SIMILARITY.
 FT DISULFID 781 803 BY SIMILARITY.
 FT DISULFID 811 853 BY SIMILARITY.
 FT DISULFID 839 864 BY SIMILARITY.
 FT DISULFID 870 915 BY SIMILARITY.
 FT DISULFID 901 926 BY SIMILARITY.
 FT DISULFID 931 973 BY SIMILARITY.
 FT DISULFID 959 984 BY SIMILARITY.
 FT DISULFID 989 1032 BY SIMILARITY.
 FT DISULFID 1018 1043 BY SIMILARITY.
 FT DISULFID 1048 1091 BY SIMILARITY.
 FT DISULFID 1077 1102 BY SIMILARITY.
 FT DISULFID 1109 1152 BY SIMILARITY.
 FT DISULFID 1138 1163 BY SIMILARITY.
 FT DISULFID 1167 1218 BY SIMILARITY.
 FT DISULFID 1201 1228 BY SIMILARITY.
 FT CARBOHYD 529 529 POTENTIAL.
 FT CARBOHYD 718 718 POTENTIAL.
 FT CARBOHYD 802 802 POTENTIAL.
 FT CARBOHYD 822 822 POTENTIAL.
 FT CARBOHYD 882 882 POTENTIAL.
 FT CARBOHYD 911 911 POTENTIAL.
 FT CARBOHYD 1029 1029 POTENTIAL.
 FT CARBOHYD 1095 1095 POTENTIAL.
 FT VARIANT 402 402
 FT CONFLICT 21 21
 FT CONFLICT 30 30
 FT CONFLICT 34 34
 FT CONFLICT 53 54
 FT STRAND 870 871
 FT STRAND 876 876
 FT TURN 877 878
 FT STRAND 879 881
 FT STRAND 890 891
 FT TURN 893 894
 FT STRAND 896 900
 FT TURN 903 904
 FT STRAND 907 907
 FT STRAND 912 916
 FT TURN 917 918
 FT STRAND 919 920
 FT STRAND 926 926
 FT SEQUENCE 1231 AA: 139125 MW: C65EC8CF8800B3FD CRC64;
 Query Match 100.0%; Score 2029; DB 1; Length 1231;
 C->Q (IN REF. 3).
 T->V (IN REF. 3).
 T->Q (IN REF. 3).
 RP->IL (IN REF. 2).
 /FTID=VAR.001979.
 H->Y.
 POTENTIAL.
 POTENTIAL.

FT	CARBOHYD	676	676	POTENTIAL.
FT	CARBOHYD	721	721	POTENTIAL.
FT	CARBOHYD	773	773	POTENTIAL.
FT	CARBOHYD	801	801	POTENTIAL.
FT	CARBOHYD	1030	1030	POTENTIAL.
FT	CARBOHYD	1061	1061	POTENTIAL.
FT	CARBOHYD	1225	1225	POTENTIAL.
SQ	SEQUENCE	1234 AA;	139082 MM;	C5AC02341B957F7 CRC64;

Query Match	70.9%;	Score 1439;	DB 1;	Length 1234;
Best Local Similarity	67.0%;	Pred. No. 0.00e+00;		
Matches 177;	Conservativity	39;	Mismatches 48;	Indels 0; Gaps 0;

Db	19	EDCKGPPRSENEIISGSWSOEYEGTQATYKCBPGYRTITGIYKCKNKNKMYASNP	SR	78
QY	1	EDCNELPPRRNTEIILTGSMSDQTYTEGTOAIYKCBPGYRSIGNVIMCRKEWALNPLR	60	
Db	79	ICRKKPCGHPGDTPFGSFRLAVGSOFEFGAKVYVYCDGQYOLLEIDYREGCAGDWINDI	138	
QY	61	KCKRPPCGHPGDTPTGFTFTLLTGSNFEYEGYKAVYTCNNGYOLLEINREDTIGWINDI	120	
Db	139	PLCEYVKKCLPYELENGRIVTSAALTDDEYYPFGQVYRPECNSGKIRGKEIHCSENGLW	198	
QY	121	PLCEYVKKCLPYAPENNGIVSAEMPDREYHFHGAVFRVCSGKIRGDEMHCSDDGEW	180	
Db	199	SNEKPRCEVETLTPPRVENGDIWVKPYKKNERYHYKCKHGYKPERGDAVCTGSGMS	258	
QY	181	SKEKPRCEVETLTPPRVENGDIWVKPYKKNERYHYKCKHGYKPERGDAVCTGSGMS	240	
Db	259	OPCEKRCSPYLLINGIYTPHRI	282	
QY	241	LPCEKRCSDNPIYIPNGDYSPLRI	264	

RESULT	3	STANDARD;	PRT;	263	AA.
ID	VCP VACCV				
AC	P10998:				
DT	01-JUL-1989 (Rel. 11, Created)				
DT	01-JUL-1989 (Rel. 11, Last sequence update)				
DT	15-JUL-1998 (Rel. 36, Last annotation update)				
DE	COMPLEMENT CONTROL PROTEIN PRECURSOR (VCP) (SECRETORY PROTEIN 35)				
GN	C3.				
OS	Vaccinia virus (strain WR), and Vaccinia virus (strain Copenhagen).				
OC	Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;				
CC	Orthopoxvirus.				
RN	[1]				
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 20-37.				
RC	STRAIN-WR.				
RX	MEDLINE: 88318974.				
RA	Kotwal G.J., Moss B.,				
RT	"Vaccinia virus encodes a secretory polypeptide structurally related				
RL	to complement control proteins.";				
RN	Nature 335:176-178(1988).				
RP	[2]				
RC	SEQUENCE FROM N.A.				
RX	STRAIN-WR.				
RA	MEDLINE: 89073756.				
RT	Kotwal G.J., Moss B.,				
RL	"Analysis of a large cluster of nonessential genes deleted from a				
RN	vaccinia virus terminal transposition mutant.";				
RP	Virology 167:524-537(1988).				
RC	[3]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-COPENHAGEN;				
RX	MEDLINE: 91021027.				
RA	Goebel S.J., Johnson G.P., Perkus M.E., Davis S.W., Winslow J.P.,				
RT	Paolletti E.;				
RL	"The complete DNA sequence of vaccinia virus.";				
RN	Virology 179:247-266(1990).				
RP	[4]				
RC	COMPLETE GENOME.				
RP	STRAIN-COPENHAGEN;				
RC					

```

RA Goebelt J., Johnson G.P., Perkus M.E., Davis S.W., Winiglow J.P.,
RA Polevitz E.;
RL virology 179:517-563(1990).
RP [5]
RN FUNCTION.
RX MEDLINE: 92115714.
RA Isaacs S.N., Kotwal G.J., Moss B.;
RT "Vaccinia virus complement-control protein prevents
RT antibody-dependent complement-enhanced neutralization of infectivity
RT and contributes to virulence."
RN Proc. Natl. Acad. Sci. U.S.A. 89:628-632(1992).
[6]
RP STRUCTURE BY NMR OF 146-263.
RX MEDLINE: 97446168.
RA Wiles A.P., Shaw G., Bright J., Perczel A., Campbell I.D.,
RA Barlow P.N.;
RT "NMR studies of a viral protein that mimics the regulators of
RT complement activation.";
RL J. Mol. Biol. 272:253-265(1997).
CC -I- FUNCTION: SERVES TO PROTECT THE VIRUS AGAINST COMPLEMENT ATTACK BY
CC INHIBITING BOTH CLASSICAL AND ALTERNATIVE PATHWAYS OF COMPLEMENT
CC ACTIVATION. BINDS C3B AND C4B.
CC -I- SIMILARITY: BELONGS TO THE SUPERFAMILY OF THE REGULATORS OF
CC COMPLEMENT ACTIVATION (RCA).
CC -I- SIMILARITY: CONTAINS 4 SUSHI (SCR) REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X13166; CAA31564.1; -.
DR EMBL: M22812; AAA69605.1; -.
DR EMBL: M35027; AAA47997.1; -.
DR PIR: A31005; MNAVSP.
DR PDB: 1VVC; 03-DEC-97.
DR PDB: 1VVD; 03-DEC-97.
DR PDB: 1VE; 03-DEC-97.
DR PFAM: PF00084; sushi; 4.
KW Signal; Repeat; Sushi; 3d-structure.
FT SIGNAL 1 19
FT CHAIN 20 263 COMPLEMENT CONTROL PROTEIN.
FT DOMAIN 20 262 4 X SUSHI (SCR) REPEATS.
FT REPEAT 20 82 SUSHI 1.
FT REPEAT 85 144 SUSHI 2.
FT REPEAT 147 202 SUSHI 3.
FT REPEAT 205 262 SUSHI 4.
FT DISULFID 21 70 BY SIMILARITY.
FT DISULFID 54 81 BY SIMILARITY.
FT DISULFID 86 126 BY SIMILARITY.
FT DISULFID 112 143 BY SIMILARITY.
FT DISULFID 148 190 BY SIMILARITY.
FT DISULFID 176 201 BY SIMILARITY.
FT DISULFID 206 248 BY SIMILARITY.
FT DISULFID 234 261 BY SIMILARITY.
SQ SEQUENCE 263 AA; 28629 MW; E4322CC9A6FE997 CRC64;

Query Match 19.1%; Score 388; DB 1; Length 263;
Best Local Similarity 34.2%; Pred. No. 7.02e-75;
Matches 80; Conservative 37; Mismatches 96; Indels 21; Gaps 17;

Db 40 ANANINIGDTLEYCLDGRGROKMGPIYACTGTGW-TL--FNDGIRRRPSPRDINDGQ 96
OY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 20 SDQYYPESTALTYICRGRYS--LGNVIMVCRGGEWALPLRKCQRRPGHGDPFGR 77
OY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 97 LDI-GG-V-DFGSSITYSCNSGYHLTGEKSYCELGSTGSMVNNPEAPICEVYKCSPPS 153
OY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
78 FTLLGNWVEFGYKAVYTCNENGYLLGEIN-YRE-CDTDS-WTNIDPLICEVYKCLPYTA 133
OY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
154 ISNR--HNGE-DF-YTDSGYVYISCNISGYSLIGNSGYLCSGGE-WS-DPTQCIIVKCP 207

```


FT DISULFID 146 176 BY SIMILARITY.
 FT DISULFID 181 223 BY SIMILARITY.
 FT DISULFID 209 240 BY SIMILARITY.
 FT DISULFID 245 287 BY SIMILARITY.
 FT DISULFID 273 299 BY SIMILARITY.
 FT DISULFID 303 343 BY SIMILARITY.
 FT DISULFID 329 355 BY SIMILARITY.
 FT DISULFID 359 400 BY SIMILARITY.
 FT DISULFID 386 413 BY SIMILARITY.
 FT CARBOHYD 74 74 POTENTIAL.
 FT CARBOHYD 227 227 POTENTIAL.
 FT CARBOHYD 275 275 POTENTIAL.
 FT CARBOHYD 292 292 POTENTIAL.
 FT CARBOHYD 366 366 POTENTIAL.
 FT CARBOHYD 381 381 POTENTIAL.
 FT CARBOHYD 428 428 POTENTIAL.
 SQ SEQUENCE 469 AA; 51551 MW; 41E137CB8DC6321 CRC64;

Query Match 16.8%; Score 341; DB 1; Length 469;
 Best Local Similarity 28.9%; Pred. No. 5,236-62;
 Matches 78; Conservative 57; Mismatches 109; Indels 26; Gaps 23;

DB 60 PPAIPALPASDVNRDFESHRTKYECLPGYGRGSRMMVYCKPGEM-ETIS-VE-CA 116
 QY 7 PPRNTEILGSMDSQ-Y-PEGTOAIKCRPGY-RSGNVIWCR-KGEWALNPLKQ 63
 DB 117 KKCRNGYLDNG-Y-VNGETI-TFGSOLFSCQGFILVSGST-SSCEVRGKGVAMSNP 172
 QY 64 KRCHGHDGDPFGFTFLTGNVFEYGVKAVYTCNEGYYQLGELINRCDD-D-G--WTND 119
 DB 173 FPCVIVKCGPPDISNGK-HSGF-E-DF-YPRNHGISTYCDGFRVGVSPFGICVYVK 228
 QY 120 IPICEVVKCPVAPENKGLVSSAMEPDRYHFGQAVFCNSGKLEGGEMHCS--DD 177
 DB 229 TVPVWSSPPTCKIKIOSPNILHGVIVSGKATYTHRDSVRLACNGTVLRGRHVLTCQ 288
 QY 178 GF--WEKREKCKVEISCKSPDIVINGSPISQ-KIIVKENEFQYKCMNGYYSERGDVCT 234
 DB 289 GNGNWSLPTCEP-DCDLPALVNGYYSM 317
 QY 235 ESG-WRPDPSCEKSCD-NPYIPNGDYSP 262

RESULT 10
 AC LEM3_RAW STANDARD; PRT: 768 AA.
 AC P98106;
 DT 01-FEB-1996 (Rel. 33 Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE P-SELECTIN PRECURSOR (GRANULE MEMBRANE PROTEIN 140) (GMP-140) (PADGM)
 DE (CD62P) (LEUKOCYTE-ENDOTHELIAL CELL ADHESION MOLECULE 3) (LECAM3).
 GN SELP.
 OS Rattus norvegicus (Rat).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 CC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LUNG;
 RX MEDLINE: 94333817.
 RA Auchampach J.A., Oliver M.G., Anderson D.C., Manning A.M.;
 RT "Cloning, sequence comparison and in vivo expression of the gene
 RT encoding rat P-selectin.";
 RL Gene 145:251-255(1994).
 CC -1- FUNCTION: CA(2+)-DEPENDENT RECEPTOR FOR MYELOID CELLS THAT BINDS
 CC TO CARBOHYDRATES ON NEUTROPHILS AND MONOCYTES. MEDIATES THE
 CC INTERACTION OF ACTIVATED ENDOTHELIAL CELLS OR PLATELETS WITH
 CC LEUKOCYTES. THE LIGAND RECOGNIZED IS STALY-LEWIS X.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN ALL TISSUES EXAMINED: SPLEEN,
 CC LUNG, BRAIN, LIVER, HEART, KIDNEY, THYMUS, SMALL INTESTINE.
 CC -1- INDUCTION: ACUTE INFLAMMATION (PROBABLY).
 CC -1- SIMILARITY: TO OTHER SELECTINS/LECAMs.
 CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.

CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
 CC -1- SIMILARITY: CONTAINS 8 SUSHI (SCR) REPEATS; RAT P-LECTIN LACKS THE
 CC HUMAN SUSHI-2 EQUIVALENT.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: L23088; AAA60325.1; -
 DR HSSP: P16109; 1FSB.
 DR PRINTS; PRO00343; SELECTIN.
 DR PROSITE; PS00022; EGF_1; 1.
 DR PROSITE; PS0186; EGF_2; 1.
 DR PROSITE; PS00615; C-TYPE LECTIN_1; 1.
 DR PROSITE; PS50041; C-TYPE LECTIN_2; 1.
 DR PFAM: PF00059; lectin_c7_1.
 DR PFAM: PF00084; sushi; 8.
 DR Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;
 KW Selectin; Signal; Sushi; Repeat.
 FT SIGNAL 1 41
 FT CHAIN 42 768
 FT DOMAIN 42 709
 FT TRANSMEM 710 733
 FT DOMAIN 734 768
 FT DOMAIN 38 158
 FT DOMAIN 159 195
 FT DOMAIN 199 700
 FT REPEAT 199 258
 FT REPEAT 261 320
 FT REPEAT 323 382
 FT REPEAT 385 444
 FT REPEAT 447 506
 FT REPEAT 509 568
 FT REPEAT 579 638
 FT REPEAT 641 700
 FT REPEAT 60 158
 FT DISULFID 131 150
 FT DISULFID 168 183
 FT DISULFID 185 194
 FT DISULFID 200 244
 FT DISULFID 230 257
 FT DISULFID 262 306
 FT DISULFID 292 319
 FT DISULFID 324 368
 FT DISULFID 354 381
 FT DISULFID 386 430
 FT DISULFID 416 443
 FT DISULFID 448 492
 FT DISULFID 478 505
 FT DISULFID 510 554
 FT DISULFID 540 567
 FT DISULFID 580 624
 FT DISULFID 610 637
 FT DISULFID 642 686
 FT DISULFID 672 699
 FT CARBOHYD 45 45
 FT CARBOHYD 54 54
 FT CARBOHYD 107 107
 FT CARBOHYD 212 212
 FT CARBOHYD 347 347
 FT CARBOHYD 456 456
 FT CARBOHYD 603 603
 FT CARBOHYD 654 654
 FT CARBOHYD 661 661
 FT CARBOHYD 679 679
 FT SITE 756 759
 SQ SEQUENCE 768 AA; 83517 MW; 26FDE8A5F3F1316 CRC64;

Query Match 16.8%; Score 341; DB 1; Length 768;

MEDLINE: 92345617.
 Sanders W.E. Jr., Wilson R.W., Ballantyne C.M., Beaudet A.L.;
 "Molecular cloning and analysis of in vivo expression of murine P-selectin."
 Blood 80:795-800(1992).
 -1- FUNCTION: CA12+-DEPENDENT RECEPTOR FOR MELOID CELLS THAT BINDS TO CARBOHYDRATES ON NEUTROPHILS AND MONOCYTES. MEDIATES THE INTERACTION OF ACTIVATED ENDOTHELIAL CELLS OR PLATELETS WITH LEUKOCYTES. THE LIGAND RECOGNIZED IS SIALYL-LEWIS X.
 -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 -1- TISSUE SPECIFICITY: STORED IN THE ALPHA-GRANULES OF PLATELETS AND WEIBEL-PALADE BODIES OF ENDOTHELIAL CELLS. UPON CELL ACTIVATION BY AGONISTS, P-SELECTIN IS TRANSPORTED RAPIDLY TO THE CELL SURFACE.
 -1- INDUCTION: BY TUMOR NECROSIS FACTOR ALPHA.
 -1- SIMILARITY: TO OTHER SELECTINS/LECTINS.
 -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
 -1- SIMILARITY: CONTAINS 8 SUSHI (SCR) REPEATS; MOUSE P-LECTIN LACKS THE HUMAN SUSHI-2 EQUIVALENT.

 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

 EMBL: M87861; AAA40008.1; -
 EMBL: M72332; AAA37712.1; -
 PIR: A42755; A42755.
 DR HSSP: P16109; 1FSB.
 DR MGD: MGI:98280; SERP.
 DR PRINTS: PRO0343; SELECTIN.
 DR PROSITE: PS00022; EGF_1; 1.
 DR PROSITE: PS01186; EGF_2; 1.
 DR PROSITE: PS00615; C-TYPE LECTIN_1; 1.
 DR PROSITE: PS50041; C-TYPE LECTIN_2; 1.
 DR PFAM: PF00008; EGF_1.
 DR PFAM: PF00009; Lectin_C; 1.
 DR PFAM: PF00084; sush1; 8.
 DR Cell adhesion; transmembrane; Glycoprotein; EGF-like domain; Lectin;
 Selectin; Signal; Sush1; Repeat.
 KW SIGNAL.
 FT CHAIN 1 41
 FT DOMAIN 42 768
 FT TRANSMEM 710 733
 FT DOMAIN 734 768
 FT DOMAIN 158 158
 FT DOMAIN 159 195
 FT DOMAIN 199 700
 FT REPEAT 199 258
 FT REPEAT 261 320
 FT REPEAT 323 382
 FT REPEAT 385 444
 FT REPEAT 447 506
 FT REPEAT 509 568
 FT REPEAT 579 638
 FT REPEAT 641 700
 FT REPEAT 700 768
 FT DISULFID 131 150
 FT DISULFID 163 174
 FT DISULFID 174 183
 FT DISULFID 185 194
 FT DISULFID 200 244
 FT DISULFID 230 257
 FT DISULFID 262 306
 FT DISULFID 292 319
 FT DISULFID 324 368
 FT DISULFID 354 381
 FT DISULFID 386 430
 FT DISULFID 416 443

FT DISULFID 448 492 BY SIMILARITY.
 FT DISULFID 478 505 BY SIMILARITY.
 FT DISULFID 510 554 BY SIMILARITY.
 FT DISULFID 540 567 BY SIMILARITY.
 FT DISULFID 580 624 BY SIMILARITY.
 FT DISULFID 610 637 BY SIMILARITY.
 FT DISULFID 642 686 BY SIMILARITY.
 FT DISULFID 672 699 BY SIMILARITY.
 FT CARBOHYD 398 398 POTENTIAL.
 FT CARBOHYD 603 603 POTENTIAL.
 FT CARBOHYD 654 654 POTENTIAL.
 FT CARBOHYD 661 661 POTENTIAL.
 FT CARBOHYD 679 679 POTENTIAL.
 FT SITE 756 759 POTENTIAL.
 FT CONFLICT 724 724 A -> E (IN REF. 2).
 SQ SEQUENCE 768 AA: 83098 MW: 5173074D2F66E8 CRC64;

 Query Match 16.7%; Score 339; DB 1; Length 768;
 Best Local Similarity 30.0%; Pred. No. 1,83e-61;
 Matches 51; Conservative 35; Mismatches 75; Indels 9; Gaps 9;
 Db 220 FEFNSQCTFSCAEYELDPGPEL-QCLASGIWNNPKCDVOCOSLEAPPHGTM-AC-M 276
 QY 86 FEYGVKAVYTCNEGQYLGEINRECDTG-WTNDIPICEVVKLPVTAPENKIVSSAM 144
 Db 277 HP1AFAVDSSCKFECPGGRARAGSNTLHCIGSGQNSPLPTCAIACEPEPIHGSD 336
 QY 145 EPDRHYHGGAVRYCNSGYKIEGDEEHGSDGFWKEREKRCVEISCKSPDV-ING-S-P 202
 Db 337 CVPSTGTFEGYNSCTFLCAEGFVLKGNDAICADSGQMTAPAPCEALOC 386
 QY 203 -ISOKIYKENERFOYKCMNGEYSEHGDAVCTESG-WR-PLRSCERKSC 249

 RESULT 13
 ID APOH_BOVIN STANDARD; PRT; 345 AA.
 AC P17690; Q28052;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE BETA-2-GLYCOPROTEIN I PRECURSOR (APOLIPOPROTEIN H) (APO-H).
 GN APOH.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae;
 OC Bovinae; Bos.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-LIVER.
 RA Gao B., Virmant M., Romm E., Lazar-Wesley E., Sakaguchi K.,
 RA Appella E., Kunos G., Takacs L.;
 RL Submitted (DEC-1992) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 4-345 FROM N.A., PARTIAL SEQUENCE, AND DISULFIDE BONDS.
 RC TISSUE-LIVER.
 RA MEDLINE: 92232647.
 RX Bendixen E., Halkier T., Magnusson S., Sottirup-Jensen L.,
 RA Kristensen T.;
 RT "Complete primary structure of bovine beta 2-glycoprotein I: localization of the disulfide bridges."
 RL Biochemistry 31:3611-3617(1992).
 RN [3]
 RP SEQUENCE OF 20-345, CARBOHYDRATE-BINDING SITES, AND DISULFIDE BONDS.
 RC TISSUE-PLASMA;
 RA MEDLINE: 92089075.
 RX Kato H., Enjyoji K.-I.;
 RT "Amino acid sequence and location of the disulfide bonds in bovine beta 2 glycoprotein I: the presence of five Sushi domains."
 RL Biochemistry 30:11687-11694(1991).
 RN [4]
 RP SEQUENCE OF 20-41.
 RX MEDLINE: 90226328.
 RA Li O., Blacher R., Esch F., Congote L.F.;

RT alternation of anticardiolipin binding.";
 RL Int. Immunol. 3:1217-1221(1991).
 RP [6]
 RA SEQUENCE FROM N.A.
 RT Kristensen T.;
 RL "Structure of the human beta-2-glycoprotein I gene."
 RT Submitted (FEB-1997) to the EMBL/Genbank/DBJ databases.
 RN [7]
 RP SEQUENCE OF 20-345, CARBOHYDRATE-BINDING SITES, AND DISULFIDE BONDS.
 RX MEDLINE: 84222015.
 RA Lozier J., Takahashi N., Putnam F.W.;
 RT "Complete amino acid sequence of human plasma beta 2-glycoprotein I.";
 RL Proc. Natl. Acad. Sci. U.S.A. 81:3640-3644(1984).
 RN [8]
 RP DISULFIDE BONDS IN C-TERMINAL DOMAIN.
 RX MEDLINE: 93050249.
 RA Steinkasserer A., Barlow P.N., Willis A.C., Kertesz Z.,
 RT Campbell I.D., Sim R.B., Norman D.G.;
 RL "Activity, disulphide mapping and structural modelling of the fifth
 domain of human beta 2-glycoprotein I.";
 RT FEBS Lett. 313:193-197(1992).
 RN [9]
 RP ANALYSIS OF CARBOHYDRATES.
 RX MEDLINE: 97293942.
 RA Gambino R., Rulu G., Pagano G., Cassader M.;
 RT "Qualitative analysis of the carbohydrate composition of
 apolipoprotein H.";
 RL J. Protein Chem. 16:205-212(1997).
 RN [10]
 RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS).
 RX MEDLINE: 99437994.
 RA Bouma B., de Groot P.G., van Den Elsen J.M.H., Ravelli R.B.G.,
 RT Schouten A., Stimmelink M.J.A., Dersken R.H.W.M., Kroon J., Gros P.;
 RL "Adhesion mechanism of human beta(2)-glycoprotein I to phospholipids
 based on its crystal structure.";
 RT EMBO J. 18:5166-5174(1999).
 RN [11]
 RP VARIANT LEU-266.
 RX MEDLINE: 93273313.
 RA Steinkasserer A., Doerner C., Wuerzner R., Sim R.B.;
 RT "Human beta 2-glycoprotein I: molecular analysis of DNA and amino
 acid polymorphism.";
 RL Hum. Genet. 91:401-402(1993).
 RN [12]
 RP VARIANT ASN-107.
 RX MEDLINE: 97369481.
 RA Sanghera D.K., Kristensen T., Hamman R.F., Kamboh M.I.;
 RT "Molecular basis of the apolipoprotein H (beta 2-glycoprotein I)
 protein polymorphism.";
 RL Hum. Genet. 100:57-62(1997).
 RN [13]
 RP VARIANTS GLY-325 AND SER-335.
 RX MEDLINE: 9721791.
 RA Sanghera D.K., Wagenknecht D.R., McIntyre J.A., Kamboh M.I.;
 RT "Identification of structural mutations in the fifth domain of
 apolipoprotein H (beta-2-glycoprotein I) which affect phospholipid
 binding.";
 RL Hum. Mol. Genet. 6:311-316(1997).
 CC -I- FUNCTION: BINDS TO VARIOUS KINDS OF NEGATIVELY CHARGED SUBSTANCES
 CC SUCH AS HEPARIN, PHOSPHOLIPIDS, AND DEXTRAN SULFATE. MAY PREVENT
 CC ACTIVATION OF THE INTRINSIC BLOOD COAGULATION CASCADE BY BINDING
 CC TO PHOSPHOLIPIDS ON THE SURFACE OF DAMAGED CELLS.
 CC -I- TISSUE SPECIFICITY: PLASMA.
 CC -I- SIMILARITY: CONTAINS 4 SUSHI (SCR) REPEATS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: X58100; CAA4113.1; -
 DR EMBL: X53595; CAA37664.1; -
 DR EMBL: X57847; CAA40977.1; -
 DR EMBL: M62839; AAA51766.1; -
 DR EMBL: S80305; AAB21330.1; -
 DR EMBL: Y11493; CAA72279.1; -
 DR EMBL: Y11494; CAA72279.1; JOINED.
 DR EMBL: Y11495; CAA72279.1; JOINED.
 DR EMBL: X53595; CAA72279.1; JOINED.
 DR EMBL: Y11496; CAA72279.1; JOINED.
 DR EMBL: Y11497; CAA72279.1; JOINED.
 DR EMBL: Y11498; CAA72279.1; JOINED.
 DR PIR: S17178; NBRU.
 DR PIR: B43286; B43286.
 DR PDB: 1QUB; 08-OCT-99.
 DR MTM: 138700; -
 DR PFM: PFM0084; sushi. 4.
 DR Heparin-binding; Glycoprotein: Plasma; Repeat: Sushi; Signal;
 KW Polymorphism: 3D-structure.
 FT SIGNAL 1 19
 FT CHAIN 20 345
 FT DOMAIN 22 261
 FT REPEAT 22 80
 FT REPEAT 83 138
 FT REPEAT 141 201
 FT REPEAT 204 261
 FT DOMAIN 263 345
 FT DISULFID 23 66
 FT DISULFID 51 79
 FT DISULFID 84 124
 FT DISULFID 110 137
 FT DISULFID 142 188
 FT DISULFID 174 200
 FT DISULFID 205 248
 FT DISULFID 234 260
 FT DISULFID 264 315
 FT DISULFID 300 325
 FT DISULFID 307 345
 FT CARBOHYD 149 149
 FT CARBOHYD 162 162
 FT CARBOHYD 183 183
 FT CARBOHYD 193 193
 FT CARBOHYD 253 253
 FT VARIANT 107 107
 FT VARIANT 266 266
 FT VARIANT 325 325
 FT VARIANT 335 335
 FT CONFLICT 121 121
 FT CONFLICT 188 188
 FT SEQUENCE 345 AA; 38298 MW; 63101704F8EDEEF3 CRC64;
 SQ
 Query Match 16.4%; Score 333; DB 1; Length 345;
 Best Local Similarity 30.1%; Pred. No. 7; 71e-60;
 Matches 72; Conservative 43; Mismatches 102; Indels 22; Gaps 18;
 Db 41 YERGERITTSCKRGYSGRGMKFCPLGLW-PINTL-KCPHYCPAGIENCAVRYT 98
 QY 24 YPGTQAIYKCRPGYSLNVLN-VCR-KGEWALNPLKCKRPGCGHDPFGFTILT 81
 Db 99 ---TFEPNTISFCNTGFLNCAISA-KTEGKWSPELPVCAPIPCPSIPFATLR 154
 QY 82 GGNVFEIGVAVYTCNMGYGLGEIYRCDIDG-WTNDIPICEVYKCLPVPAPENGRKI- 139
 Db 155 VYKPSAGNSNLYR-DTAV-PECLPQHAMGNDITITTTGNNTKL-PEGREVKCFPSRP 211
 QY 140 V-SSAMEPDREYHFGAVRVCSNGYKIBGEDMHCSDDGFWSKERPKCVETISCKSPD-V 197

Fri Jun 9 10:53:38 2000

US-09-316-163-10.rsp

Page 16

Db 212 DNGFVNYPAKPTLYYKDKATFG--CHDGYSLDGPEIECTKXIGNMSAMPSCKA-SCKVP 267
OY 198 ING--S-PISOKIITYKENRFRQYKCNMGIEYSERGDVAVCTESG-WRPLPSCHEKSCDNP 252

Search completed: Thu Jun 8 21:42:12 2000
Job time : 13 secs.

Release 3.1A John F. Collins, Biocomputing Research Unit
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

```
Run on: Thu Jun 8 21:42:31 2000; MasPar time 26.07 Seconds
704.848 Million cell updates/sec
```

Tabular output not generated.

Title: >US-09-316-163-10
Description: (1-265) from US09316163.pep

Sequence: 1 EDCNLEPPRRNTEILTGSWS.....EKSCDNPIPNGDYSPLRK 265

Scoring table: PAM 150

Searched: 225878 seqs, 69334122 residues

```
Post-processing: Minimum Match 08
                  Listing first 45 summaries
```

Database:

1:sp_archaea 2:sp_bacteria 3:sp_fungi 4:sp_human
 5:sp_invertebrate 6:sp_mammal 7:sp_mmc 8:sp_organalle
 9:sp_phase 10:sp_plant 11:sp_rodent 12:sp_unclassified
 13:sp_vertebrate 14:sp_virus

Statistics: Mean 43.031; Variance 62.408; scale 0.690

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	2029	100.0	449	4	014570	COMPLEMENT FACTOR H PR	0.00e+00
2	952	46.9	669	6	028085	CCP MODULES 3-12, WITH	4.00e-23
3	587	28.9	1053	13	091275	COMPLEMENT REGULATOR	7.83e-12
4	436	21.5	645	14	09FRU2	COMPLEMENT BINDING PRO	2.63e-85
5	422	20.8	550	14	040912	ORE 04.	1.61e-81
6	421	20.7	550	14	P88903	ORE 4.	3.00e-81
7	398	19.6	259	14	P87616	41KBP FRAGMENT FROM LE	4.65e-75
8	396	19.5	657	4	014006	COMPLEMENT H FACTOR (F	1.60e-74
9	382	18.8	360	14	09YR08	COMPLEMENT CONTROL PRO	8.91e-71
10	357	17.6	263	14	089859	HOMOLOG OF VACCINIA VI	3.96e-64
11	357	17.6	263	14	007033	D12L PROTEIN.	3.96e-64
12	355	17.5	263	14	089076	B19L.	1.34e-63
13	353	17.4	560	5	Q22328	CSMD12 T07H6.	4.52e-63
14	351	17.3	533	11	008569	ALLOSOMAL MATRIX COMPO	1.35e-62
15	350	17.2	263	14	089061	D15L.	2.81e-62
16	341	16.8	559	11	063135	COMPLEMENT REGULATOR	6.61e-60
17	338	16.7	349	4	015429	CD46.	4.07e-59
18	337	16.6	417	11	035570	512 ANTIGEN (FRAGMENT)	7.44e-59
19	336	16.6	740	4	095508	D7180M13.1.2 (SELECTIN	1.36e-58
20	334	16.5	974	5	P91658	FORWARDED.	4.56e-58

21	331	16.3	285	6	019126	MEMBRANE COFACTOR PROT	2.79e-57
22	331	16.3	285	6	028290	MEMBRANE COFACTOR PROT	2.79e-57
23	331	16.3	754	6	P79138	CELL ADHESION MOLECULE	2.79e-57
24	328	16.2	369	6	P79138	MEMBRANE COFACTOR PROT	1.70e-56
25	328	16.2	2039	4	016745	COMPLEMENT RECEPTOR 1	1.70e-56
26	328	16.2	016744	4	016744	COMPLEMENT RECEPTOR 1	1.70e-56
27	326	16.1	2489	11	060736	ZONA PELLUCIDA 3 RECEPTOR	5.68e-56
28	326	16.1	579	11	060736	MEMBRANE COFACTOR PROT	3.46e-55
29	322	15.9	285	6	019121	COMPLEMENT RECEPTOR 1	6.30e-55
30	322	15.9	661	6	029531	COMPLEMENT RECEPTOR 1	6.30e-55
31	320	15.8	2014	6	029530	COMPLEMENT RECEPTOR 1	2.10e-54
32	318	15.7	522	6	028769	MEMBRANE COFACTOR PROT	6.96e-54
33	316	15.6	377	6	062838	MEMBRANE COFACTOR PROT	1.31e-53
34	316	15.6	363	6	002839	PORCINE MEMBRANE COFACTOR	2.27e-53
35	316	15.6	378	6	062837	MEMBRANE COFACTOR PROT	2.31e-53
36	316	15.6	497	11	063612	512 ANTIGEN PRECURSOR	1.39e-52
37	313	15.4	1911	6	029552	COMPLEMENT RECEPTOR 1	1.39e-52
38	308	15.2	482	6	028982	E-SELECTIN	2.76e-51
39	300	14.8	833	5	099254	FACTOR B SPeF	2.76e-51
40	299	14.7	679	11	099254	COMPLEMENT RECEPTOR TV	3.24e-51
41	299	14.7	740	4	095507	COMPLEMENT FACTOR H-RE	5.88e-49
42	297	14.6	315	6	028770	COMPLEMENT RECEPTOR (F	1.93e-48
43	295	14.5	347	4	P78361	DECAV-ACCELEATION FAC	6.31e-44
44	292	14.4	522	6	028797	UNKNOWN PROTEIN (P8AGM	4.73e-44
45	291	14.3	646	6	029097	P-SELECTIN PRECURSOR	6.74e-44

ALIGNMENTS

Query Match	100.0%	Score 2029;	DB 4;	Length 449;
-------------	--------	-------------	-------	-------------

Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 265; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 19 EDCEMLPRRMTTELTGSMQSYTEGTOAIYKCPGRSLGNVIMVCRKGEWALNPLR 78
QY 1 EDCEMLPRRMTTELTGSMQSYTEGTOAIYKCPGRSLGNVIMVCRKGEWALNPLR 60
Db 79 KQKRPCHGPDTEFGFTLLTGNVFEYGVAVYTCNGYOLGELINRECDTGWTNDI 138
QY 61 KQKRPCHGPDTEFGFTLLTGNVFEYGVAVYTCNGYOLGELINRECDTGWTNDI 120
Db 139 PICVAVCLPTAPENKIVSSAMEPDEHYHFGQAVRFVNSGKIEGDEMHCSDDGF 198
QY 121 PICVAVCLPTAPENKIVSSAMEPDEHYHFGQAVRFVNSGKIEGDEMHCSDDGF 180
Db 199 SKERPKCVESICKSPDINGSISQKIIYKNERFQYKCNMGYSERGDVCTESGMRP 258
QY 181 SKERPKCVESICKSPDINGSISQKIIYKNERFQYKCNMGYSERGDVCTESGMRP 240
Db 259 LPSCSEKSCNPIYIPNDYSPRLIK 283
QY 241 LPSCSEKSCNPIYIPNDYSPRLIK 265

RESULT 2
ID 028085 PRELIMINARY; PRT; 669 AA.
AC 028085;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, last sequence update)
DT 01-NOV-1996 (TREMBlrel. 12, last annotation update)
DE CCP MODULES 3-12, WITH PARTS OF CCP 2 AND 13 (FRAGMENT).
OS BCS taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia;
OC Euteria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae;
OC Bovinae; Bos.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE; 96202005.
RA SOAMES C.J., DAY A.J., SIM R.B.;
RT "Prediction from sequence comparisons of residues of factor H involved
in the interaction with complement component C3b."
RL Biochem. J. 315:523-531(1996).
DR HSSP; P10998; IYVC.
DR PFAM; PF00084; sushi; 11.
FT NON_TER 1
FT NON_TER 669
SQ SEQUENCE 669 AA; 75683 MW; FAF0D174 CRC32;

Query Match 46.9%; Score 952; DB 6; Length 669;
Best Local Similarity 64.3%; Pred. No. 4.00e-230;
Matches 117; Conservative 29; Mismatches 36; Indels 0; Gaps 0;

Db 7 LAEGQFEYGAQVYTCDEGYMGEANFRCDTNGMTNDIPICVAVCLPTBENGKI 66
QY 80 LTGQVFEYGAQVYTCDEGYMGEANFRCDTNGMTNDIPICVAVCLPTBENGKI 139
Db 67 FSDALEPDEYTGQVYFECNSGYMLDGPQIHCSSAGWSAEKPKVEIFCKRPVILN 126
QY 140 VSSAMEPDEHYHFGQAVRFVNSGKIEGDEMHCSDDGFMSKPKVEISCKSPDVIN 199
Db 127 GQAVLPKATYKQNEVQYRCAGFEYGGRGDTVCTKSGWTPAPTCIETCDPPRIPNGVY 186
QY 200 GSPISQKIIYKNERFQYKCNMGYSERGDVCTESGMRPLPSCSEKSCNPIYIPNDY 259
Db 187 RP 188
QY 260 SP 261

RESULT 3
ID 091275 PRELIMINARY; PRT; 1053 AA.

AC 091275;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, last sequence update)
DT 01-NOV-1996 (TREMBlrel. 12, last annotation update)
DE COMPLEMENT REGULATORY PLASMA PROTEIN.
OS Paratubax nebulifer (barred sand bass).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
OC Neopterygii; Teleostei; Euteleostei; Acanthopterygii; Percomorpha;
OC Perciformes; Percoidae; Serranidae; Paratubax.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE; 94318039.
RA DAMEY A., KAIJOH T., ZIPPEL P.F., GIGLI I.;
RT "Cloning and characterization of a cDNA representing a putative
complement-regulatory plasma protein from barred sand bass (Paratubax
nebulifer)."
RL Biochem. J. 301:391-397(1994).
DR EMBL; L21703; AAA82556.1; -.
DR HSSP; P08603; IHEH.
DR PFAM; PF00084; sushi; 16.
SQ SEQUENCE 1053 AA; 117597 MW; 0DF68EDB CRC32;

Query Match 28.9%; Score 587; DB 13; Length 1053;
Best Local Similarity 37.1%; Pred. No. 7.83e-127;
Matches 91; Conservative 41; Mismatches 101; Indels 12; Gaps 9;

Db 47 EASYPEGROVRCNGVY-S-GFKLVCEGKWTGCA--KCOPRSCGHGDAOFADPHL 102
QY 21 DQYPEGTOAIYKCRFGYSLGNVIMVCRKGEWALNPLRKQKCRGCHGPDTEFTL 80
Db 103 AEGNDVFSGKYVYTCQKQYQVSRNTNRYCAEGKGVVPCESQOC-PLIHVDNNVQV 161
QY 81 TGNVFEYGVAVYTCNEGYOLGELINRECDTGWTNDIPICVAVCLPTAPENGKI 140
Db 162 IGG-PE-EATFGNVVRFSCRSREILDSPLYCDEMGSPVKKCAICAIPIEN 218
QY 141 SSAMEPDEHYHFGQAVRFVNSGKIEGDEMHCSDDGFMSKPKVEISCKSPDVIN 199
Db 219 GNVPGAIKREKNDVLEHCDNAFKHIDRPSTCIKQIKAEWSPPLGESIKRCLTMDG 278
QY 200 GSPISQKIIYKNERFQYKCNMGYSERGDV--CTESGMRPLPSCSEKSCNPIYIPN 256
Db 279 TRYEP 283
QY 257 GDYSP 261

RESULT 4
ID 09WRU2 PRELIMINARY; PRT; 645 AA.
AC 09WRU2;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, last sequence update)
DT 01-NOV-1999 (TREMBlrel. 12, last annotation update)
DE COMPLEMENT BINDING PROTEIN.
OS Macaca mulatta rhadinovirus 17577.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Rhadinovirus.
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE; 99174001.
RA SEARLES R.P., BERGUAM E.P., AXTHELM M.K., WONG S.W.;
RT "Sequence and genomic analysis of a Rhesus macaque rhadinovirus with
similarity to Kaposi's sarcoma-associated herpesvirus/human
herpesvirus 8."
RL J. Virol. 73:3040-3053(1999).
DR EMBL; AF083501; AAD21332.1; -.
SQ SEQUENCE 645 AA; 71526 MW; 60FB82D6 CRC32;

Query Match 21.5%; Score 436; DB 14; Length 645;
Best Local Similarity 30.9%; Pred. No. 2.63e-85;
Matches 72; Conservative 50; Mismatches 94; Indels 17; Gaps 14;

```

Db 337 EKVSGASVELICRPGFTKQSTVSECLSNGTWAPNA--KCHAKKCP7POELNGEYI 394
QY 22 QTPGEGQAIYKCRPGSRISGNVIMV-C-RKGEVALNPLKRCOKRCPGHPGTPTGT 79
Db 395 VTSGEDAFKGTITTYKCNNGCYOLLGSMVRIICMLKODLKVDPEAPAPICDICKCPKPPQ 454
QY 80 LTGG-NVFEYGVAVYTCNNGCYOLLG-E-IN-Y--REC-DTIDGNTINDIPICEVVKCLPYTA 133
Db 455 ITNGKY--HPEV-DE-YOYLDVTFTSCNRPDSIVGDEMTICISNT-WMKPFPCEQITGS 509
QY 134 PENKTIYSAMEPREHVFQAVAFVNCNGYKIEGDEMHCSDDGFWSKPKVCVEISCK 193
Db 510 APNIAHKKLLTGSSSVYKRGQSVTIGETGPTLIGSEIISCKSSMPPLPTC 562
QY 194 SPDVINGSPIS-OKIYKENERFOYKCMNGEYSEKGDVCTESGM-RPLPSC 244

RESULT 5
ID 040912 PRELIMINARY: PRT: 550 AA.
AC 040912:
DT 01-JAN-1998 (Tremblrel. 05, Created)
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)
DT 01-NOV-1999 (Tremblrel. 12, Last annotation update)
DE ORF 04
OS Kaposi's sarcoma-associated herpesvirus (KSHV) (Human herpesvirus 8).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Rhadinovirus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 97296220.
RA NEPEL F., ALBRECHT J.C., FLECKENSTEIN B.;
"Cell-homologous genes in the Kaposi's sarcoma-associated rhadinovirus
human herpesvirus 8: determinants of its pathogenicity?";
RL J. Virol. 71:4187-4192(1997).
DR EMBL: U93872; AAB62602.1; -.
DR HSSP: P10998; IYVC.
DR PFAM: PF00084; sush1. 4.
SQ SEQUENCE 550 AA; 60648 MW; 558089B4 CRC32;

Query Match 20.8%; Score 422; DB 14; Length 550;
Best Local Similarity 31.9%; Pred. No. 1,61e-81;
Matches 75; Conservative 46; Mismatches 95; Indels 19; Gaps 15;

Db 52 RCRSGYTYARNITATCLOGGTW-S-EPTATCNKKSCEPNGEIONGKVIFFHGODALXYG 109
QY 33 KCRPGYRSLG-NVIMVC-RKGEVALNPLKRCOKRCPGHPGTPTGTTLTG-NVFEYX 89
Db 110 ANISYVCNDEGYFLVGREYRYCMIGASGOMAMSSPPFCEKEK-H-R-P---KIENGDF 163
QY 90 VKAVYTCNNGCYOLLG-E-INRECDTDG--WTNDIPICEVVKCLPYTAPENGKIVSSAM 144
Db 164 KPKDYEVNDVAHFECNNGCYTLVGHSTACAVANNWTSMPTCELACGKPSVTHGYPI 223
QY 145 EPREYH-FEQAVAFVNCNGYKIEGDEMHCSDDGFWSKPKVCVEISCKSPVINGSPI 203
Db 224 QGSFLTKHKOSYTFACNDGFVLGSPITTCNTVTEMDPLPKVCLEDDIDDPNNSN 278
QY 204 SQ-KIYKENERFOYKCMNGEYSEKGDVCTESGM-RPLPSCSEKSCDNPYIPN 256

RESULT 6
ID P88903 PRELIMINARY: PRT: 550 AA.
AC P88903:
DT 01-MAY-1997 (Tremblrel. 03, Created)
DT 01-MAY-1997 (Tremblrel. 03, Last sequence update)
DT 01-NOV-1999 (Tremblrel. 12, Last annotation update)
DE ORF 4.
OS Kaposi's sarcoma-associated herpesvirus (KSHV) (Human herpesvirus 8).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Rhadinovirus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 97121460.

```

```

RA RUSSO J.J., BOHENZKY R.A., CHIEN M.C., CHEN J., YAN M., MADDALENA D.,
RA PARRY J.P., PERIZZI D., EDELMAN I.S., CHANG Y., MOORE P.S.;
RT "Nucleotide sequence of the Kaposi sarcoma-associated herpesvirus
RT (HHV8).";
RT Proc. Natl. Acad. Sci. U.S.A. 93:14862-14867(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE: 97094384.
RA MOORE P.S., BASHOFF C., WEISS R.A., CHANG Y.;
RT "Molecular mimicry of human cytokine and cytokine response pathway
RT genes by KSHV.";
RL Science 274:1739-1744(1996).
DR EMBL: U75698; AAC57082.1; -.
DR HSSP: P10998; IYVC.
DR PFAM: PF00084; sush1. 4.
SQ SEQUENCE 550 AA; 60687 MW; CEA2CAEC CRC32;

Query Match 20.7%; Score 421; DB 14; Length 550;
Best Local Similarity 31.9%; Pred. No. 3.00e-81;
Matches 75; Conservative 46; Mismatches 95; Indels 19; Gaps 15;

Db 52 RCRSGYTYARNITATCLOGGTW-S-EPTATCNKKSCEPNGEIONGKVIFFHGODALXYG 109
QY 33 KCRPGYRSLG-NVIMVC-RKGEVALNPLKRCOKRCPGHPGTPTGTTLTG-NVFEYX 89
Db 110 ANISYVCNDEGYFLVGREYRYCMIGASGOMAMSSPPFCEKEK-H-R-P---KIENGDF 163
QY 90 VKAVYTCNNGCYOLLG-E-INRECDTDG--WTNDIPICEVVKCLPYTAPENGKIVSSAM 144
Db 164 KPKDYEVNDVAHFECNNGCYTLVGHSTACAVANNWTSMPTCELACGKPSVTHGYPI 223
QY 145 EPREYH-FEQAVAFVNCNGYKIEGDEMHCSDDGFWSKPKVCVEISCKSPVINGSPI 203
Db 224 QGSFLTKHKOSYTFACNDGFVLGSPITTCNTVTEMDPLPKVCLEDDIDDPNNSN 278
QY 204 SQ-KIYKENERFOYKCMNGEYSEKGDVCTESGM-RPLPSCSEKSCDNPYIPN 256

RESULT 7
ID P87616 PRELIMINARY: PRT: 259 AA.
AC P87616:
DT 01-MAY-1997 (Tremblrel. 03, Created)
DT 01-MAY-1997 (Tremblrel. 03, Last sequence update)
DT 01-NOV-1999 (Tremblrel. 12, Last annotation update)
DE 41KB FRAGMENT FROM LEFT END OF GENOME.
GN D17L OR C17L.
OS Cowpox virus (CPV).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-GRI-90;
RA MEDLINE: 97068532.
RA SAFRONOV P.F., PETROV N.A., RIAZANKINA O.I., TOTMENIN A.V.,
RA SHCHELKUNOV S.N., SANDAKHCHIEV L.S.;
RT "Genes of a circle of hosts for the cowpox virus.";
RL Dokl. Akad. Nauk 349:829-833(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN-GRI-90;
RA MEDLINE: 98229462.
RA SHCHELKUNOV S.N., SAFRONOV P.F., TOTMENIN A.V., PETROV N.A.,
RA RIAZANKINA O.I., GUNOROV V.V., KOTVAL G.J.;
RT "The genomic sequence analysis of the left and right species-specific
RT terminal region of a cowpox virus strain reveals unique sequences and
RT a cluster of intact ORFs for immunomodulatory and host range
RT proteins.";
RL Virology 243:432-460(1998).
DR EMBL: X94355; CA64102.1; -.
DR EMBL: Y11842; CA672567.1; -.
DR HSSP: P10998; IYVC.
DR PFAM: PF00084; sush1. 4.
SQ SEQUENCE 259 AA; 28193 MW; 76531F63 CRC32;

```


SEQUENCE 263 AA: 28816 MW: C9DFBA61 CRC32:

Query Match 17.6%: Score 357; DB 14; Length 263;

Best Local Similarity 33.6%: Pred. No. 3,96e-64; Mismatches 95; Indels 23; Gaps 19;

Matches 79; Conservative 38; Mismatches 95; Indels 23; Gaps 19;

Db 40 ANANYNIGDTIEYLCPLGPKRQKMGPIYAKCTGTG-TL-FNQCIRKRCPSPRDINDGH 96
 20 SDQTPETGTAIYKCRPGYRS--LGNVIMYCRKGEWALNPLKRCQKRCGHPGDPFPFGT 77
 78 FTLLGNGVFEYGVKAVYTCNCGYOLLGEIN-Y-RECDTDG--WTNDIPICEVVKC-LPVT 132
 Db 97 LDI-GG-V-DEGSSITTCNSGYLLIGEKSYCKLGSTGSMWNPAPICSEYKCOLPPS 153
 78 FTLLGNGVFEYGVKAVYTCNCGYOLLGEIN-Y-RECDTDG--WTNDIPICEVVKC-LPVT 132
 Db 154 IS-NGR--HNGYN-DF-YTDGSVVTCNSGYLLIGNSGYLCSGGE-WSNP-PTCQIYVC 206
 133 APENKTIYSSAMEPDRHYFGQAVRFVNCNSGYKIEGDEMHCSDDGFWSKPKVCVEISC 192
 133 APENKTIYSSAMEPDRHYFGQAVRFVNCNSGYKIEGDEMHCSDDGFWSKPKVCVEISC 192
 Db 207 PHPTLLNGYLLSSGFKRSYSYNDVDFCTKYGYKLGSSSSSTCSPGNTWQPELPKC 261
 193 KSPDVINGSPISQ-KIYKENERFOYKCNMGYEYSEKGDVCTESG-WRP-LPSC 244
 193 KSPDVINGSPISQ-KIYKENERFOYKCNMGYEYSEKGDVCTESG-WRP-LPSC 244

RESULT 11
 ID 007033 PRELIMINARY; PRT; 263 AA.

AC 007033:
 DT 01-NOV-1996 (TREMblrel. 01, Created)

DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)

DT 01-NOV-1999 (TREMblrel. 12, Last annotation update)

DE D12L PROTEIN.

GN D12L.

OS Variola virus.

OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;

OC Orthopoxvirus.

OC [1]

RP SEQUENCE FROM N.A.

RC STRAIN-INDIA-1967, SSP. MAJOR;

RA SHCHELKUNOV S.N., TOTMENIN A.V.;

RA SHCHELKUNOV S.N., BLINOV V.M., SANDAKHCHIEV L.S.;

RA OLENIINA L.V., CHIRIKOVA G.B., SANDAKHCHIEV L.S.;

RT "Analysis of the nucleotide sequence of 53 kbp from the right terminus

RT of the genome of variola major virus strain India-1967.";

RL Virus Res. 34:207-236(1994).

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN-INDIA-1967, SSP. MAJOR;

RA SHCHELKUNOV S.N., TOTMENIN A.V.;

RA "Two types of deletions in orthopoxvirus genomes.";

RL Virus Genes 9:231-245(1995).

DR EMBL: X69198; CAA48953.1.;

DR PIR: C36838; C36838.

DR HSSP: P10998; IYVC.

DR PFAM: PF00084; sush1.4.

DR SEQUENCE 263 AA: 28789 MW: 6436C44C CRC32:

Query Match 17.6%: Score 357; DB 14; Length 263;

Best Local Similarity 33.6%: Pred. No. 3,96e-64; Mismatches 95; Indels 23; Gaps 19;

Matches 79; Conservative 38; Mismatches 95; Indels 23; Gaps 19;

Db 40 ANANYNIGDTIEYLCPLGPKRQKMGPIYAKCTGTG-TL-FNQCIRKRCPSPRDINDGH 96
 20 SDQTPETGTAIYKCRPGYRS--LGNVIMYCRKGEWALNPLKRCQKRCGHPGDPFPFGT 77
 97 LDI-GG-V-DEGSSITTCNSGYLLIGEKSYCKLGSTGSMWNPAPICSEYKCOLPPS 153

78 FTLLGNGVFEYGVKAVYTCNCGYOLLGEIN-Y-RECDTDG--WTNDIPICEVVKC-LPVT 132

Db 154 IS-NGR--HNGYN-DF-YTDGSVVTCNSGYLLIGNSGYLCSGGE-WSNP-PTCQIYVC 206

133 APENKTIYSSAMEPDRHYFGQAVRFVNCNSGYKIEGDEMHCSDDGFWSKPKVCVEISC 192

207 PHPTLLNGYLLSSGFKRSYSYNDVDFCTKYGYKLGSSSSSTCSPGNTWQPELPKC 261

193 KSPDVINGSPISQ-KIYKENERFOYKCNMGYEYSEKGDVCTESG-WRP-LPSC 244

RESULT 12
 ID 089076 PRELIMINARY; PRT; 263 AA.

AC 089076:
 DT 01-NOV-1996 (TREMblrel. 01, Created)

DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)

DT 01-NOV-1999 (TREMblrel. 12, Last annotation update)

DE B19L.

OS Variola virus.

OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;

OC Orthopoxvirus.

OC [1]

RP SEQUENCE FROM N.A.

RC STRAIN-GARCIA-1966;

RA MASSUNG R.F., LOPAREV V.N., KNIGHT J.C., CHIZHIKOV V.E., PARSONS J.M.;

RA TOTMENIN A.V., SHCHELKUNOV S.N., ESPOSITO J.J.;

RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.

DR EMBL: U18338; AAA69358.1.;

DR HSSP: P10998; IYVC.

DR PFAM: PF00084; sush1.4.

DR SEQUENCE 263 AA: 28842 MW: D66A9287 CRC32:

Query Match 17.6%: Score 355; DB 14; Length 263;

Best Local Similarity 33.6%: Pred. No. 1.34e-63;

Matches 79; Conservative 38; Mismatches 95; Indels 23; Gaps 19;

Db 40 ANANYNIGDTIEYLCPLGPKRQKMGPIYAKCTGTG-TL-FNQCIRKRCPSPRDINDGH 96

20 SDQTPETGTAIYKCRPGYRS--LGNVIMYCRKGEWALNPLKRCQKRCGHPGDPFPFGT 77

97 LDI-GG-V-DEGSSITTCNSGYLLIGEKSYCKLGSTGSMWNPAPICSEYKCOLPPS 153

78 FTLLGNGVFEYGVKAVYTCNCGYOLLGEIN-Y-RECDTDG--WTNDIPICEVVKC-LPVT 132

Db 154 IS-NGR--HNGYN-DF-YTDGSVVTCNSGYLLIGNSGYLCSGGE-WSNP-PTCQIYVC 206

133 APENKTIYSSAMEPDRHYFGQAVRFVNCNSGYKIEGDEMHCSDDGFWSKPKVCVEISC 192

207 PHPTLLNGYLLSSGFKRSYSYNDVDFCTKYGYKLGSSSSSTCSPGNTWQPELPKC 261

193 KSPDVINGSPISQ-KIYKENERFOYKCNMGYEYSEKGDVCTESG-WRP-LPSC 244

RESULT 13

ID 022328 PRELIMINARY; PRT; 560 AA.

AC 022328:
 DT 01-NOV-1996 (TREMblrel. 01, Created)

DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)

DT 01-NOV-1999 (TREMblrel. 12, Last annotation update)

DE COSMID T07H6.

GN T07H6.5.

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditidae;

OC Rhabditina; Rhabditioidea; Rhabditidae; Pelodetidae; Caenorhabditis.

OC [1]

RP SEQUENCE FROM N.A.

RC STRAIN-BRISTOL N2;

RA WILSON R., AINSOUGH R., ANDERSON K., BAYNES C., BERKS M.,

RA BONFIELD J., BURTON J., CONNELL M., CORSEY T., COOPER J., COULSON A.,

RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAYELLO A., FULTON L.,

RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,

RA JONES M., KESHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
 RA LIGHTNING J., LOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
 RA PASONS J., PERCY C., RIFKEN L., ROOPEA A., SAUNDERS D., SHOWNKEEN R.,
 RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., STOLSTON J.,
 RA THERREY-MEGE J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLIDMAN P.,
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RT elegans." ;
 RT Nature 368:32-38(1994) .
 RN [2]
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA GEISEL C. ;
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
 RP [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA WATERSTON R. ;
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U53344; AAA96225.1; -
 DR HSSP: P10998; IYVC.
 DR PFAM: PF00084; sushi; 7.
 SQ SEQUENCE 560 AA; 61619 MW; 3A10AF63 CRC32;

Query Match	17.4%;	Score 353;	DB 5;	Length 560;
Best Local Similarity	32.2%;	Pred. No. 4.52e-63;		
Matches	68;	Conservative	42;	Mismatches 82;
			Indels 19;	Gaps 15;

```

D6 117 AQWGPB-LR-CKNRACPDDEDIENG---JREGOTFEPHHVKSCNPGFLVSTGTS-RQ 170
QY 51 GEWALNMPRLRCQMRPGGHODTFFGFTITLGGVNEYCYKAVYTCNEGJQJLGEIYRE 110
D6 171 CSNNGEWTNEPANCKATECSRSPSSPLHGKVVGSLL-F--YO-S-VYTYSCDHGRVLVQ 224
QY 111 CDDTG-WTNDIPLEEVVCKLPVPAPENGKIYSSAMEBPDREIHHGQANRAFCNSGYKIED 169
D6 225 VQRICLAEGTNGGNEPRCEIIRCSVLPTLNGYITIEGSETSFGAVAVFRCLTMTHE-GA 282
QY 170 EEMHCSDDGWSKRPCKVCBEISCKS-PDVYINGSPISCKIYKENERFOYKCMNGYEYSR 228
D6 283 SKAKMEDGOWSAPIRCLA-SCVPHIONG 312
QY 229 GDAVCTESSG-WR-PLBSCIEKSCDNPIPIPNG 257

```

ID	RESULT_14		PRELIMINARY:	PRT:	533 AA.
AC	008569;				
DT	01-JUL-1997	(TREMBLrel. 04,			
DT	01-JUL-1997	(TREMBLrel. 04,			
DT	01-NOV-1999	(TREMBLrel. 12,			
OS	ACROSOMAL MATRIX COMPONENT AM67	PRECUSOR.			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;				
OC	Eutheria; Rodentia; Hystriocognathi; Cavidae; Cavia.				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-HARTLEY; TISSUE-TESTIS;				
RX	MEDLINE: 97284752.				
RA	FOSTER J A., FRIDAY B.B., MAULT M.T., BIOBELI C., WINFREY V.P.,				
RA	OLSON G.E., KIM K.S., GEPTON G.L.;				
RT	"AM67, a secretory component of the guinea pig sperm acrosomal matrix				
RT	is related to mouse sperm protein sp56 and the complement component 4				
RT	binding proteins.";				
RL	J. Biol. Chem. 272:12714-12722(1997).				
DR	Emr.; U75654; AAC13888.1; -.				
DR	HSSP; P10998; IVVC.				
DR	PFAM; PF00084; sush1; 7.				
KM	SIGNAL.				
FT	CHAIN	1	28	POTENTIAL.	
FT	CHAIN	29	533	ACROSOMAL MATRIX COMPONENT AM67.	
SEQUENCE	533 AA; 59772 MW; 60DEA526 CRC32;				

Query Match	17.3%	Score 351;	DB 11;	Length 533;
Best Local Similarity	27.0%	Pred. No. 1,53e-62;		
Matches	67;	Conservative	66;	Mismatches 93; Indels 22; Gaps 16;
Db	47	ENREFGTALAKNCHRGVYRWNSSHVVICDINGSMT-YNVF-CARRKRCRNGELANGKVE	103	
QY	21	DDTYEGEGQAIYKCCPPGRSGNINMVC-RKGEVWAIPLFKCCKRCBGHGDPFSTFT	79	
Db	104	IIT-DLL-FGSTIEHSCSGKSYLIGSTF-SQCESQAGKTVPMDSDLPEPCVYIKQSPDIS	160	
QY	80	LTGQWVFEXYGVKAVYTCNEGYYQLLGEINRYRCDIDG---WTNDIPICEVYKCLPYAPE	135	
Db	161	NGK-HSGN-DDDL-VTYSLTYVCDPRYSILGNASISCLVANKTVGWSNPPTECKVT	217	
QY	136	NGKIYSSAMEDREHREHQAVFCNCSNGYKIEGDEEHNC--SDD-GFWSKERKCKVEIS	191	
Db	218	CRQPHIKRGIPLSGFGFYTYTKDLIVLSCKKGYLLRDSISIIHCNANSKVPSPITCPCNG	277	
QY	192	CKSPDIVNGSPISQ-KIITYKENERFYCKMGMGYLSRGRDANC-TESGWRP-LPSCCKES	248	
Db	278	CIDLPPEVP	285	
QY	249	C-DNPYIP	255	

RESULT	15	PRELIMINARY;	PRT;	263 AA.
ID	089061			
AC	089061;			
DT	01-NOV-1996 (Tremblrel.. 01, Created)			
DT	01-NOV-1996 (Tremblrel.. 01, Last sequence update)			
DT	01-NOV-1999 (Tremblrel.. 12, Last annotation update)			
DE	D15L.			
OS	Varicella virus.			
OC	Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;			
OC	Orthopoxvirus.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-CONGO-1965;			
RA	MATSUME R.F., LOPAREV V.N., KNIGHT J.C., CHIZHIKOV V.E., PARSONS J.M.			
RA	TOSHEKIN A.V., SHEKELKUNOV S.N., ESPOSITO J.J.			
RL	Submitted (DEC-1994) to the EMBL/Genbank/DBD databases.			
DR	EMBL: U18337; AAA69317.1, -			
DR	HSSP: P10998; 1YVC.			
DR	PRAM: PF00084; sushi: 4.			
SQ	SEQUENCE 263 AA; 28843 MW; 55866165 CRC32;			

[illegible]

Search completed: Thu Jun 8 21:43:01 2000
Job time : 30 secs.

 W39154
 (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

MPerch_p protein - protein database search, using Smith-Waterman algorithm
 Run on: Thu Jun 8 21:44:51 2000; Maspar time 11.39 Seconds
 684.453 Million cell updates/sec
 Tabular output not generated.

Title: >US-09-316-163-11
 Description: (1-329) from US09316163.pep
 Perfect Score: 2539
 Sequence: 1 EDCNELPRRNTTEILTGSMS.....PDIKHGGLYHEHMRPFPPV 329

Scoring table: PAM 150
 Gap 11

Searched: 188963 segs, 23686106 residues

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database: a-geneseq36
 I-geneseqp

Statistics: Mean 33.142; Variance 124.795; scale 0.266

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	1832	72.2	240	1	Human partial Complement	3.40e-190
2	1550	65.0	216	1	Clone PRB9F410 CFH r	1.47e-169
3	452	17.8	578	1	Human complement factor	9.41e-33
4	433	17.1	581	1	Human C4 binding prote	1.05e-33
5	409	16.1	290	1	Human partial compleme	3.96e-31
6	409	16.1	778	1	Human partial compleme	3.96e-31
7	409	16.1	1930	1	Amino acid sequence of	3.96e-31
8	406	16.1	2317	1	Human complement recep	3.96e-31
9	406	16.0	1537	1	Partial human compleme	8.29e-31
10	406	16.0	2039	1	Human complement type	1.36e-30
11	404	15.9	543	1	CR1-4 (85R, 87N) analo	1.74e-30
12	403	15.9	543	1	CR1-4 (95R, 103E) anal	1.74e-30
13	402	15.8	543	1	CR1-4 (114S) analogue.	2.22e-30
14	400	15.8	543	1	CR1-4 (64K) analogue.	3.62e-30
15	389	15.7	543	1	CR1-4 (78T, 79D) analo	4.65e-30
16	389	15.7	543	1	CR1-4 (369-376 SRKPPIC	4.65e-30
17	389	15.7	543	1	CR1-4 (52S, 55S, 54P)	4.65e-30
18	399	15.7	543	1	CR1-4 (347, 367 NAH) a	4.65e-30
19	399	15.7	543	1	CR1-4 (347, 349Y) ana	4.65e-30
20	398	15.7	543	1	CR1-4 (109N, 110A, 111	5.95e-30
21	398	15.7	543	1	CR1-4 (92T) analogue.	5.95e-30
22	397	15.6	543	1	CR1-4 (121Q) analogue.	7.61e-30
23	395	15.6	543	1	CR1-4 (318-321 RNP) a	1.25e-29

ID	Score	Query Match	Length	DB ID	Description	Pred. No.
24	395	15.6	543	1	R28549	1.25e-29
25	395	15.6	543	1	R28563	1.25e-29
26	394	15.5	543	1	R28561	1.59e-29
27	394	15.5	543	1	R28551	1.59e-29
28	393	15.5	543	1	R28545	2.04e-29
29	393	15.5	543	1	R28566	2.04e-29
30	393	15.5	543	1	R28570	2.04e-29
31	393	15.5	543	1	R28562	2.04e-29
32	393	15.5	543	1	CR1-4 (116K) analogue.	3.33e-29
33	390	15.4	543	1	CR1-4 (35E) analogue.	4.26e-29
34	391	15.4	543	1	CR1-4 (57V, 59K) analo	3.33e-29
35	389	15.3	543	1	R28544	5.44e-29
36	387	15.2	543	1	R28556	8.90e-29
37	386	15.2	543	1	CR1-4 (94H) analogue.	1.14e-28
38	385	15.2	543	1	CR1-4 (92T, 94H) analo	1.45e-28
39	385	15.2	543	1	CR1-4 (35E, 37Y) analo	1.45e-28
40	382	15.0	543	1	CR1-4 (114-117SRK, 12	3.04e-28
41	382	15.0	577	1	CR1-4 (44T, 47D, 49Y)	3.04e-28
42	382	15.0	577	1	Membrane co-factor pro	1.32e-27
43	376	14.8	830	1	P-selectin.	4.50e-27
44	371	14.6	263	1	Deduced sequence of co	4.06e-26
45	362	14.3	579	1	Amino acid sequence of	6.77e-24
	341	13.4	302	1	Herpesvirus saimiri SC	

ALIGNMENTS

RESULT 1
 ID W39154 standard; Protein; 240 AA.

AC W39154:
 DT 27-APR-1998 (first entry)
 DE Human partial Complement factor H protein fragment 1.
 KW Complement factor H; tumour associated antigen; renal cancer;
 KM urogenital cancer; medicament; modulator.
 OS Homo sapiens.
 PN W09738136-A1.
 PD 16-OCT-1997
 PE 09-APR-1997; US-05710
 PF 06-MAR-1997; US-812481.
 PR 09-APR-1996; US-015083.
 PR 09-APR-1996; US-630048.
 PR 06-MAR-1997; US-038614.
 PA (BARD-) BARD DIAGNOSTIC SCI INC.
 PI Enfield DL, Hass GM, Kinders RJ;
 DR WPI: 97-51742/47.
 DR N-PSDB: V02790.

PT Treating or screening for cancer, e.g. renal or urogenital cancer -
 PT by modulating or detecting tumour associated human complement factor
 PT H related antigen, or nucleic acid encoding it
 PS Example 6B: Fig 6B; 104pp; English.
 CC This partial protein sequence represents a region of the human
 CC tumour-associated complement factor H (CFH). This sequence is used
 CC in the identification of complement factor H related proteins and
 CC antigens isolated from clone PRB9F410 (see W39155). The detection of
 CC such proteins and a CFH antigens can be used in screening or for the
 CC treatment of renal or urogenital cancer, e.g. bladder, cervical or
 CC prostate cancer. Agents that may modulate this antigen could be used in
 CC the manufacture of a medicament for the treatment of a tumour cell.
 SQ Sequence 240 AA:

Query Match 72.2%; Score 1832; DB 1; Length 240;
 Best Local Similarity 100.0%; Pred. No. 3.40e-190;
 Matches 240; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ID	Score	Query Match	Length	DB ID	Description	Pred. No.
Db	1	FTLTGNGVEYGVKAVYTCNENGYQLIGETINRCDTGTNDIPICEVYKCLPYAPENG	60			
Qy	78	FTLTGNGVEYGVKAVYTCNENGYQLIGETINRCDTGTNDIPICEVYKCLPYAPENG	137			
Db	61	KYSSAEPDREYHFGQAVFVNSGYKTEGDEHMSDGFESKPKPCVETISCKSPDY	120			
Qy	138	KYSSAEPDREYHFGQAVFVNSGYKTEGDEHMSDGFESKPKPCVETISCKSPDY	197			
Db	121	INGSPISOKIITYENENFYCKNMGEYSERGDVAVTESGWRPLPSCEKSCDNPIYING	180			

OY 158 INGSPIKOKIYKENEREFOYKCNMGYSESGDAVCTESGMRPLPSCSEKSCDNPIYNG 257
 DB 181 DYSPLRIKRTGDEITTYOORNGEYPTATRGNTAKCTSTGWIAPAPRCLTKCDYDPIKHGGL 240
 OY 258 DYSPLRIKRTGDEITTYOORNGEYPTATRGNTAKCTSTGWIAPAPRCLTKCDYDPIKHGGL 317
 RESULT 2
 ID W39155 standard; Protein; 216 AA.
 AC W39155;
 DT 27-APR-1998 (first entry)
 DE Clone PR99FH410 CFH related protein fragment.
 KW Complement factor H; tumour associated antigen; renal cancer;
 KW urogenital cancer; medicament; modulator.
 OS Synthetic.
 PS WO9738136-A1.
 PF 16-OCT-1997.
 PR 09-APR-1997; U05710.
 PR 06-MAR-1997; US-812481.
 PR 09-APR-1996; US-015083.
 PR 06-MAR-1997; US-038614.
 PA (BARD-) BARD DIAGNOSTIC SCI INC.
 PI Entfeld DL, Hass GM, Kinders RJ;
 DR WPI: 97-512742/47.
 DR N-PSDB: V02791.
 PT Treating or screening for cancer, e.g. renal or urogenital cancer -
 PT by modulating or detecting tumour associated human complement factor
 PS Example 6B; Fig 6B; 104pp; English.
 CC This partial protein is found in clone PR99FH410 and represents a
 CC human tumour-associated complement protein with homology to a region of the
 CC protein and a CFH antigen can be used in screening or for the treatment
 CC of renal or urogenital cancer, e.g. bladder, cervical or prostate cancer.
 CC Agents that may modulate this antigen could be used in the manufacture of
 CC a medicament for the treatment of a tumour cell.
 SQ Sequence 216 AA;
 Query Match 65.0%; Score 1650; DB 1; Length 216;
 Best Local Similarity 99.5%; Pred. No. 1.47e-169;
 Matches 215; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 DB 1 TCNGYOLLGELINRECDTGTGNDIPICEVYKCLPTAPANGKIYSSAMPEDEYHGG 60
 OY 95 TCNGYOLLGELINRECDTGTGNDIPICEVYKCLPTAPANGKIYSSAMPEDEYHGG 154
 DB 61 AVRFVNSGKRIEGDEDMHCSDDGFMGKPKCYEISCKSPDVINGSPISOKIYKENER 120
 OY 155 AVRFVNSGKRIEGDEDMHCSDDGFMGKPKCYEISCKSPDVINGSPISOKIYKENER 214
 DB 121 FOYKCNMGYSESGDAVCTESGMRPLPSCSEKSCDNPIYNGDYSPLRIKHTGDEITY 180
 OY 215 FOYKCNMGYSESGDAVCTESGMRPLPSCSEKSCDNPIYNGDYSPLRIKHTGDEITY 274
 DB 181 QCRNGEYPTATRGNTAKCTSTGWIAPAPRCLTKCDYDPIKHGGL 216
 OY 275 QCRNGEYPTATRGNTAKCTSTGWIAPAPRCLTKCDYDPIKHGGL 310
 RESULT 3
 ID Y09065 standard; Protein; 578 AA.
 AC Y09065;
 DT 06-JUL-1999 (first entry)
 DE Human complement factor H homolog protein.
 KW Human complement factor H; immunological mechanism; complement reaction;
 KW gene therapy; immune stimulation; haematopoiesis regulation; chemotactic;
 KW tissue growth activity; anti-inflammatory; tumour inhibition;
 KW secretory signal.
 OS Homo sapiens.
 PN WO9918200-A1.
 PF 15-APR-1999.
 PF 02-OCT-1998; J04448.

PR 06-OCT-1997; JP-272837.
 PA (PROT-) PROTEGENE INC.
 PA (SAGA) SAGAMI CHEM RES CENT.
 PI Kato S, Sekine S;
 DR WPI: 99-264019/22.
 DR N-PSDB: X34737.
 PT Human proteins with secretory signal sequences and nucleotide
 PT sequences, useful in control of proliferation and differentiation of
 PT cells
 PS Claims 1; Page 55-58; 71pp; English.
 CC This DNA encodes a protein having homology to human complement factor H,
 CC which plays a role in the immunological mechanism involving the
 CC complement reaction. The protein can also be used as an antigen for
 CC preparing antibodies against the protein. The cDNA can be used as a probe
 CC for gene diagnosis and the gene for gene therapy, as well as for large-
 CC scale expression of the protein. The protein may also have immune
 CC stimulating or suppressing activity, haematopoiesis regulating activity,
 CC tissue growth activity, activity/inhibit activity, anti-inflammatory
 CC activity, tumour inhibition activity, chemotactic/chemokinetic activity,
 CC receptor/ligand activity, etc. The protein is identified by the presence
 CC of a hydrophobic N-terminal secretory signal region, knowledge of the
 CC cloning.
 SQ Sequence 578 AA;
 Query Match 17.8%; Score 452; DB 1; Length 578;
 Best Local Similarity 32.8%; Pred. No. 9.41e-36;
 Matches 82; Conservative 56; Mismatches 88; Indels 24; Gaps 20;
 DB 54 YYCDQNFVTPSGSYWDYIHCITDQMSPTVP-C-LRTCSKSDVEIENG-FIS---ESSSIY 107
 OY 94 YTCNEGYOLL-GEI-NYRECDDTGTNDIPICEVYKCL-PTAPANGKIYSSAMPEDEYHGG 150
 DB 108 ILNEETQYNKPKGYATADNGSSGISTLQNG-WSTQ-PICIKF-CDMPVENSRAKSNGM 164
 OY 151 HFGQAVRFVNSG-KIEGDEE--MHCSDGFMGKPKCYEISCKSPDVINGSPISOKI 207
 DB 165 WFKLHDITDEYCYGSESGNTTDSIVCGEDGWSHLPTCYNSSESGCPPPISNGTTS 224
 OY 208 IYKNEHNFQYKCNMGYSESGDAVCTESGMRPLPSC-EER-SCDNPIY-IPNGYSP 261
 DB 225 FPOKVYLPMSRVEYQCS-YTELQSKRYTCSNGSDSEPPRCISMKPCPEPIQGHLYY 283
 OY 262 LRIRK-HRTGDEITYOORNGEYPTATRGNTAKCTSTGWIAPAPRCLTKCDYDPIKHGGL 319
 DB 284 ENMRPPTFPV 293
 OY 320 ENMRPPTFPV 329
 RESULT 4
 ID R13490 standard; Protein; 581 AA.
 AC R13490;
 DT 30-OCT-1991 (first entry)
 DE Human C4 binding protein.
 KW C4bp; monomer; complement protein; pI0D.C4bp.3; SCR;
 KW short consensus repeat.
 OS Homo sapiens.
 FH Key
 FT peptide
 FT 1..32
 FT /label= signal_peptide
 FT 33..581
 FT /label= C4bp
 FT /label= C4dp
 FT 33..93
 FT /label= SCR8
 FT 94..155
 FT /label= SCR7
 FT 156..219
 FT /label= SCR6
 FT 220..279
 FT /label= SCR5
 FT 280..345
 FT /label= SCR4

PI Annenkov A, Chernajovsky Y;
 DR WPI: 98-568350/48.
 PT Fragment of soluble human complement receptor 1 - useful for
 PT treating T-cell or B-cell mediated immune responses e.g.
 PT inflammatory responses such as rheumatoid arthritis
 PS Discloure: Fig 1; 54pp; English.
 CC This is an amino acid sequence of the human soluble complement
 CC receptor 1 (SCR1), useful in the treatment of T-cell or B-cell
 CC mediated immune responses. It is used to inhibit a T-cell or
 CC B-cell-mediated immune response to prevent immune response-mediated
 CC tissue rejection and destruction or clearance or inactivation of an
 CC expressed protein especially from cells that have been treated by gene
 CC therapy to express the protein, e.g. dystrophin. The protein can also
 CC be used to inhibit a T-cell or B-cell-mediated inflammatory response,
 CC an interferon-gamma secretory response, autoimmune response or
 CC neurological response, e.g. Alzheimer's or Parkinson's disease or
 CC multiple sclerosis. Also the protein can be used to treat systemic
 CC lupus erythematosus, rheumatoid arthritis, myasthenia gravis,
 CC epidermis bulosa or Hashimoto's disease.
 SQ Sequence 778 AA;

Query Match 16.1%; Score 409; DB 1; Length 778;
 Best Local Similarity 31.4%; Pred. No. 3,966-31;
 Matches 97; Conservative 66; Mismatches 110; Indels 36; Gaps 30;

Db 68 PPIGTLYNECRPGYSGRPFST-ICLNKSVWTGAKD-R-CRRKSCRNPDPYNGMVAHYK 124
 QY 24 YPEGQAIYKCRPGYRSGYVIMWC-RKGEWVALNPLRCKCRPGHGDPTFTLTG 82
 Db 125 G-I-QFGSQTLYCTKGRYLIGSSSATCIISGDTYVINDNETPICRIICGLPPTIT-NGD 181
 QY 83 GNVEYGVKAVYTCNEGYQLLGE--INRYEC-DIDGWTNDIPICEVVC-LPTVAPENGK 138
 Db 182 FISTNRE-N--FHYGSVVTYRCNPGSGRKFELVGEPSIYCTSNDDQVIGMSGAPQCI 238
 QY 139 IVSSAMPEDEHYHGQAVFYVCSGY--KI-E--GDEMHK-S-DD--GFWSKEKPCY 188
 Db 239 IPNCKTPPVNENGLIVSDNLSFLNEVVERCQPGFVMKPRRYKQALNKKEPELPSC 298
 QY 189 EIS-CKSPDYINGSPISO-KIYKENRFQYCKMGEYSERGDVCTE-SGWRP-LPSC 244
 Db 299 S-RVCOPEPVLAHERQKDNFSPQGEVYFSGEPG-Y-DLGAASMRCTPGQDMSPPA 355
 QY 245 EKSCKDNP-YIPNGDYSPL-RIKHRTGDEITYCRNGEYPATRGNTA-KCTSTG-WTPA- 299
 Db 356 PTCGVKSCD 364
 QY 300 PCTLKPCD 308

RESULT 7
 ID W45899 standard; peptide; 1930 AA.
 AC W45899;
 DT 30-JUN-1998 (first entry)
 DE Human complement receptor 1 (residues 1-1929).
 KW Membrane binding element; thrombotic disease; soluble protein;
 KW complement-related disease; integral membrane protein; inflammation;
 KW short consensus repeat; SCR 1-3; CRI; complement receptor type 1.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT 1930 "Disulphide linked to Cys in peptide given
 FT /note= in W45889"
 PN W09802454-A2.
 PD 22-JAN-1998.
 PE 08-JUL-1997; E03715.
 PR 15-JUL-1996; GB-014871.
 PA (ADPR-) ADPROTECH PLC.
 PI Dodd I, Mossakowska DEI, Smith RAG;
 DR WPI: 98-110524/10.
 PT membrane binding groups - useful for treating complement-related
 PT thrombotic diseases, providing improved localisation at cellular

PT membranes
 PS Claim 22; Pages 60-61; 75pp; English.
 CC This sequence represents human complement receptor 1 (CRI, CD 35)
 CC N-terminal fragment. The invention relates to a soluble derivative (A)
 CC of a soluble polypeptide (I), which comprises at least 2 heterologous
 CC membrane-binding elements (MBE) of low membrane affinity covalently
 CC associated with (I). MBE interact, independently and with thermodynamic
 CC additivity, with components of cellular or artificial membranes exposed
 CC to extracellular fluids. (A) are used to treat disorders treatable with
 CC (I) itself, specifically inflammation or any other complement-related
 CC disorder (e.g. neurological disease, graft rejection, myocardial
 CC infarction, sepsis, rheumatoid arthritis and many others; including
 CC application to indwelling devices) and thrombolytic disease, but also to
 CC treat allergy, induce weight loss, to treat ischemia or asthma and as
 CC immuno-modulators for treating multiple sclerosis. (A) are administered
 CC orally, topically, by injection or inhalation at 0.01-10 (preferably
 CC 0.1-10) mg/kg/day.
 SQ Sequence 1930 AA;

Query Match 16.1%; Score 409; DB 1; Length 1930;
 Best Local Similarity 31.4%; Pred. No. 3,966-31;
 Matches 97; Conservative 66; Mismatches 110; Indels 36; Gaps 30;

Db 22 PPIGTLYNECRPGYSGRPFST-ICLNKSVWTGAKD-R-CRRKSCRNPDPYNGMVAHYK 78
 QY 24 YPEGQAIYKCRPGYRSGYVIMWC-RKGEWVALNPLRCKCRPGHGDPTFTLTG 82
 Db 79 G-I-QFGSQTLYCTKGRYLIGSSSATCIISGDTYVINDNETPICRIICGLPPTIT-NGD 135
 QY 83 GNVEYGVKAVYTCNEGYQLLGE--INRYEC-DIDGWTNDIPICEVVC-LPTVAPENGK 138
 Db 136 FISTNRE-N--FHYGSVVTYRCNPGSGRKFELVGEPSIYCTSNDDQVIGMSGAPQCI 192
 QY 139 IVSSAMPEDEHYHGQAVFYVCSGY--KI-E--GDEMHK-S-DD--GFWSKEKPCY 188
 Db 193 IPNCKTPPVNENGLIVSDNLSFLNEVVERCQPGFVMKPRRYKQALNKKEPELPSC 252
 QY 189 EIS-CKSPDYINGSPISO-KIYKENRFQYCKMGEYSERGDVCTE-SGWRP-LPSC 244
 Db 253 S-RVCOPEPVLAHERQKDNFSPQGEVYFSGEPG-Y-DLGAASMRCTPGQDMSPPA 309
 QY 245 EKSCKDNP-YIPNGDYSPL-RIKHRTGDEITYCRNGEYPATRGNTA-KCTSTG-WTPA- 299
 Db 310 PTCGVKSCD 318
 QY 300 PCTLKPCD 308

RESULT 8
 ID P92219 standard; protein; 2317 AA.
 AC P92219;
 DT 22-FEB-1990 (first entry)
 DE CRI protein
 KW Complement; cofactor.
 OS Homo sapiens (human).
 FH Key Location/Qualifiers
 FT 10.50
 FT /label= signal_peptide
 PN W08909220-A.
 PD 05-OCT-1989.
 PE 31-MAR-1989; U01358.
 PR 01-APR-1988; US-176532.
 PA (TCEU) T Cell Sciences Inc; (WVJO) The Johns Hopkins University;
 PA (BRIG*) The Brigham and Women's Hospital.
 PI Fearon DT, Klickstein LB, Wong W, Carson G, Concino MF, Makrides SC;
 DR WPI: 89-309498/42.
 DR N-PSDB: N91477.
 PT New nucleic acid sequences encoding new CRI protein - and its fragment,
 PT for diagnosis and control of complement-related immune defects,
 PT inflammation, myocardial infarct, etc
 PS Claim 1; fig. 1; 191pp; English.
 CC This is full-length CRI protein, and shortened forms are new, lacking
 CC the transmembrane region. The proteins and fragments bind C3b and/or

CC C4b, have cofactor activity and inhibit C3 and C5 convertase activity.
 CC In the sequence, x-untranslated region. This has 7 short consensus
 CC repeats (CR) of 4 long direct homologous repeats (LHR) (see fig. 10 in
 CC patent). Deletion mutants, eg LHR-A for binding of C4b, and LHR-C
 CC for C3b binding, be made, and LHR-B and -C can detect CR2 sequences.
 CC They are useful in diagnosing and treating immune disorders, and prevent
 CC perfusion injury.
 CC Sequence 2317 AA:

Query Match 16.1%; Score 409; DB 1; Length 2317;
 Best Local Similarity 31.4%; Pred. No. 3,96e-31;

Matches 97; Conservative 66; Mismatches 110; Indels 36; Gaps 30;

Db 72 FPIGTLYNECRPGYSGPFSI-ICLNKSVTGAKD-R-CRRKSCRRPPDVNGMVAHIK 128
 QY 24 YPEGQAIYKCRPGYSGPFSI-ICLNKSVTGAKD-R-CRRKSCRRPPDVNGMVAHIK 82
 Db 129 G-I-OFGSQIKYSGTGRILGSSATCIISGDTVMNDENPDCRIPGILPTIT-NGD 185
 QY 83 GNVFEGYKAVYTCNEGYQLGE--INRRC-DIDGWTNDIPICGVK-LPVTAPENGK 138
 Db 186 FISTNRE-N-FHGSVYTYRCNPGSGRKYVELVGEPSITCTSDNDQVIGSGPAPOCI 242
 QY 139 IYSSAMEPDREYHFGQAVRVCNSGY--KI-E--GDEMH-C-S-DD--GFMSKEKPKCV 188
 Db 243 IPNKTTPRNVENGILVSDNRLSFLSNEVEFRCPQGFVMAKGRYKQALNKKWEPELPSC 302
 QY 189 EIS-CKSPDVINGSPISQ-KIYKENEROYKCMNGIYSESGDAVCTE-SGNRP-LPSC 244
 Db 303 S-RVCCPPPDVLAHARTQRDKDNFSPGOEVEFYSCPEG-Y-DLGAASMRCTPGQDMSPPA 359
 QY 245 EKSQDNP-YIPNGDYSPL-RIKHRTGDELTTCGRNGFYPATRGNTA-KCTSTG-WIPA- 299
 Db 360 PCEVAKSCD 368
 QY 300 PRCITLKPCD 308

RESULT 9

ID R11982 standard; Protein: 1537 AA.

AC R11982;

DE 25-JUN-1991 (first entry)

KW Partial human complement type 1 receptor.

KW Complement system; C3b/C4b receptor; CRI; allergic reaction;

OS Homo sapiens.

OS Homo sapiens.

OS Homo sapiens.

OS Homo sapiens.

OS Homo sapiens.

OS Homo sapiens.

OS Homo sapiens.

OS Homo sapiens.

OS Homo sapiens.

OS Homo sapiens.

OS Homo sapiens.

OS Homo sapiens.

OS Homo sapiens.

OS Homo sapiens.

OS Homo sapiens.

OS Homo sapiens.

OS Homo sapiens.

OS Homo sapiens.

OS Homo sapiens.

OS Homo sapiens.

OS Homo sapiens.

OS Homo sapiens.

OS Homo sapiens.

OS Homo sapiens.

OS Homo sapiens.

OS Homo sapiens.

OS Homo sapiens.

OS Homo sapiens.

OS Homo sapiens.

OS Homo sapiens.

OS Homo sapiens.

PT fragments - for treatment of immune disorders, myocardial infarct,
 PT damage due to inflammation and in treatment of thrombosis
 PS Disclosure: Fig 5: 234pp: English.
 CC This sequence comprises three of the four tandem, direct, long
 CC homologous repeats of the full-length F allele of CRI. LHR-A is
 CC absent. Each LHR might represent a single C3b/C4b binding domain,
 CC making the receptor multivalent. The LHR's are composed of 7 short
 CC consensus repeats of 60-70 residues resembling the SCR's of other
 CC C3/C4 binding proteins. The protein and fragments of it having C3b
 CC and/or C4b binding activity can be used to treat immune disorders
 CC or disorders involving inappropriate complement activity.
 CC See also Q11642.
 CC Sequence 1537 AA:

Query Match 16.0%; Score 406; DB 1; Length 1537;
 Best Local Similarity 28.8%; Pred. No. 8,29e-31;

Matches 89; Conservative 70; Mismatches 114; Indels 36; Gaps 26;

Db 914 FVPGTSLNTECRPGY--FEKMFISCLNVLWSSVED--NCRKSCGPPPE-PFNGMVAHI 968
 QY 24 YPEGQAIYKCRPGYSGPFSI-ICLNKSVTGAKD-R-CRRKSCRRPPDVNGMVAHIK 82
 Db 969 NTOT-QGSTVANSNCEGRLGSPSTTGLVSGNNVTMDKAPICEIISCPEPTISNGD 1027
 QY 82 GNVFEGYKAVYTCNEGYQLGE--INRRC-DIDGWTNDIPICGVK-LPVTAPENGK 138
 Db 1028 FYSNNR-TS--FHNGTVVYVYOCHTGPDGDLVELVGERISYCTSKDDQVGVSSPPPCI 1084
 QY 139 IYSSAMEPDREYHFGQAVRVCNSG-----YKIEGDEMH-C-S-DD--GFMSKEKPKCV 188
 Db 1085 STNCTAPEVNAIRVPGNRSFSLTEIIRFCQPPFVAVGSTITVCCQNGRMPPLPHC 1144
 QY 189 EIS-CKSPDVINGSPISQ-KIYKENEROYKCMNGIYSESGDAVCTE-SGNRP-LPSC 244
 Db 1145 S-RVCCPPPDVLAHARTQRDKDNFSPGOEVEFYSCPEG-Y-DLGAASMRCTPGQDMSPPA 1201
 QY 245 EKSQDNP-YIPNGDYSPL-RIKHRTGDELTTCGRNGFYPATRGNTA-KCTSTG-WIPA- 299
 Db 1202 PRCITLKPCD 1210
 QY 300 PRCITLKPCD 308

RESULT 10

ID R11810 standard; Protein: 2039 AA.

AC R11810;

DE 25-JUN-1991 (first entry)

KW Human complement type 1 receptor.

KW Complement system; C3b/C4b receptor; CRI; allergic reaction;

OS Homo sapiens.

OS Homo sapiens.

OS Homo sapiens.

OS Homo sapiens.

OS Homo sapiens.

OS Homo sapiens.

OS Homo sapiens.

OS Homo sapiens.

OS Homo sapiens.

OS Homo sapiens.

OS Homo sapiens.

OS Homo sapiens.

OS Homo sapiens.

OS Homo sapiens.

OS Homo sapiens.

OS Homo sapiens.

OS Homo sapiens.

OS Homo sapiens.

OS Homo sapiens.

OS Homo sapiens.

OS Homo sapiens.

OS Homo sapiens.

OS Homo sapiens.

OS Homo sapiens.

OS Homo sapiens.

OS Homo sapiens.

OS Homo sapiens.

(W.I.)

Distribution rights by Oxford Molecular Ltd

Tabular output not generated.

Sequence:

Scoring table:

Searched:

2

Database:

Statistics:

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	2539	100.0	449	1	NBHHHS	complement factor H p	0.00e+00
2	2539	100.0	1231	1	NBHHH	complement factor H p	0.00e+00
3	1772	69.8	1234	1	NBMBS	complement factor H p	0.00e+00
4	1236	48.7	669	2	S65551	factor H - bovine (lr	3.71e-28
5	603	23.7	1053	2	S46199	probable complement r	2.98e-11
6	447	17.6	597	1	S53711	C4B alpha chain prec	5.01e-80
7	438	17.3	452	2	A35068	complement factor H-r	8.46e-78
8	433	17.1	303	2	H35068	apolipoprotein H-rela	1.45e-76
9	433	17.1	397	1	NBHUC4	C4b-binding protein a	4.36e-74
10	423	16.7	2014	1	I36936	complement receptor 1	1.18e-70
11	409	16.1	2039	2	A28507	complement C3b/C4d re	1.18e-70
12	409	16.1	2489	2	I73012	complement C3b/C4d re	1.98e-65
13	404	15.9	558	2	S57953	C4BP protein alpha ch	3.49e-66
14	403	15.9	661	2	I36937	complement receptor 1	6.13e-66
15	402	15.8	465	1	NBMSC4	C4b-binding protein a	1.62e-66
16	388	15.3	263	1	MWVZSP	apolipoprotein H homo	1.62e-66
17	388	15.3	560	2	I16833	hypothetical protein	2.84e-66
18	387	15.2	482	2	A34924	complement C3b/C4d re	1.35e-66
19	376	14.8	830	2	A30359	P-selectin precursor	7.21e-66
20	373	14.7	676	2	A45900	complement C3b recept	2.20e-66
21	371	14.6	345	1	NEMS	apolipoprotein H prec	2.05e-66
22	367	14.5	768	2	I35821	P-selectin - rat	3.33e-55
23	362	14.3	579	2	A56740	sperm-egg recognition	

24	363	14.3	668	2	A46013	coagulation factor XI	1.91e-59
25	360	14.2	808	1	D35069	complement factor H-r	1.01e-58
26	357	14.1	263	2	C53658	complement control pr	5.37e-58
27	357	14.1	610	1	I46001	C4b-binding protein A	5.37e-58
28	343	13.5	768	2	A42755	P-selectin precursor	1.26e-54
29	341	13.4	302	1	WMBE1E	secretory complement	3.79e-54
30	341	13.4	360	1	WMBE2E	membrane-bound comple	3.79e-54
31	340	13.4	661	1	KRFU13	coagulation factor XI	6.59e-54
32	338	13.3	349	2	G02913	spem CD46 - human (f	1.99e-53
33	338	13.3	369	2	I57998	membrane cofactor pro	1.98e-53
34	334	13.2	345	1	NBBO	apolipoprotein H prec	1.80e-52
35	335	13.2	377	2	I54479	membrane cofactor pro	1.04e-52
36	335	13.2	384	2	SC0195	membrane cofactor pro	1.04e-52
37	336	13.2	497	2	UCJ054	complement regulatory	5.98e-53
38	331	13.0	345	1	NBRU	apolipoprotein H prec	9.37e-52
39	328	12.9	362	2	UCJ194	membrane cofactor pro	4.87e-51
40	328	12.9	369	2	UCJ138	membrane cofactor pro	2.35e-50
41	325	12.8	1091	1	PL0009	complement C3d/Eps1e	3.91e-49
42	320	12.6	610	2	A43046	endothelial leukocyte	3.91e-49
43	318	12.5	610	2	A33606	endothelial leukocyte	1.17e-48
44	314	12.4	612	2	S23174	endothelial leukocyte	1.04e-47
45	309	12.2	551	2	I46709	endothelial leukocyte	1.58e-46

ALIGNMENTS

```

RESULT      1      #type complete
ENTRY      NBH0HS      complement factor H precursor, short splice form - human
TITLE      complement factor H-related protein; complement protein H
ALTERNATE_NAMES      #formal_name Homo sapiens #common_name man
ORGANISM      31-dec-1993 #sequence_revision 23-Feb-1996 #text_change
DATE      22-Jun-1999
ACCESSIONS      S03013; B60238; A27877; A61103; A26505; S10479
REFERENCE      S00254
#authors      Ripoché, J.; Day, A.J.; Harris, T.J.R.; SLM, R.B.
#journal      Biochem. J. (1988) 249:593-602
#title      The complete amino acid sequence of human complement factor
              H.
#cross-references MUID:88134059
#accession      S03013
#molecule_type mRNA
#residues      1-449 ##label RIP
#cross-references EMBL:X07523; EMBL:Y00716; NID:g32492;
              PIDN:CA030403.1; PID:g756073
#note      part of this sequence, including the amino end of the
              mature protein was confirmed by protein sequencing
              402-Tyr was also found
#note      A60238
#authors      Estaller, C.; Schwaebler, W.; Dierich, M.; Weiss, E.H.
#journal      Eur. J. Immunol. (1991) 21:799-802
#title      Human complement factor H: two factor H proteins are derived
              from alternatively spliced transcripts.
#cross-references MUID:91184292
#accession      B60238
#status      not compared with conceptual translation
#molecule_type mRNA
#residues      1-33;434-449 ##label EST
#note      Only portions of this 1.8 kilobase mRNA were sequenced
REFERENCE      A27877
#authors      Schultz, T.F.; Schwaebler, W.; Stanley, K.K.; Weiss, E.;
              Dierich, M.P.
#journal      Eur. J. Immunol. (1986) 16:1351-1355
#title      Human complement factor H: isolation of cDNA clones and
              partial cDNA sequence of the 38-kDa tryptic fragment
              containing the binding site for C3b.
#cross-references MUID:87054207
#accession      A27877
#molecule_type mRNA
#residues      1-11;55-401; 'Y', 403-449 ##label SCH
#note      an additional nucleotide present within the codon for
              glu-310 was thought to be a cloning artifact and was
              ignored in translation

```

```

REFERENCE
#authors Schwabbe, W.; Zwirner, J.; Schulz, T.F.; Linke, R.P.;
#journal Dietrich, M.P.; Weiss, E.H.
#title Eur. J. Immunol. (1987) 17:1485-1489
#cross-references Human complement factor H: expression of an additional
#accession truncated gene product of 43 kDa in human liver.
#status A61103
#molecule-type not compared with conceptual translation
#residues 27-76 ##label SC2
#note this is a partial sequence of an alternatively spliced
1.8 kilobase mRNA that is translated to yield a 43 k
form related to factor H

REFERENCE
#authors A26505
#journal Sim, R.B.; Discipio, R.G.
#title Biochem. J. (1982) 205:285-293
#cross-references Purification and structural studies on the complement-system
#accession control protein beta-1-H (factor H).
A26505

#molecule-type protein
#residues 19-20, 'Q', 22-29, 'V', 31-33, 'Q', 35 ##label SIM
A4451

REFERENCE
#authors Barlow, P.N.; Norman, D.G.; Steinkasserer, A.; Horne, T.J.;
#journal Pearce, J.; Driscoll, P.C.; Sim, R.B.; Campbell, I.D.
#title Biochemistry (1992) 31:3626-3634
#cross-references Solution structure of the fifth repeat of factor H: A second
#accession example of the complement control protein module.
MU1D:92232649
#contents annotation; NMR structure determination, residues 264-292
S10479

REFERENCE
#authors Kristensen, T.; Wetzel, R.A.; Tack, B.F.
#journal J. Immunol. (1986) 136:3407-3411
#title Structural analysis of human complement protein H: homology
with C4b binding protein, beta(2)-glycoprotein I, and the
Ba fragment of B.
#cross-references MU1D:86169701
#accession S10479

#molecule-type mRNA
#residues 226-T01,'Y',403-449 ##label KRI
#cross-references GB:M12383; NID:9180472; PIDW:AAA52013.1; PID:9180473
COMMENT Factor H has also been found bound to cell membranes in an unknown
manner. However, it has at least one cell attachment site motif
in repeat 4.
COMMENT Alternative transcripts of 4.3, 1.8, and 1.4 kilobases are expressed
in liver. See also PIR:NBH0H.

GENETICS
#gene HFI
#cross-references GDB:120041; OMIM:134370
#map_position 1q32-1q32
GENETICS
#gene GDB:HF1; HF
#cross-references GDB:129095
#map_position 1q32-1q32
#note the correspondence between the two loci and the sequences
indicated is unclear; factor H has been reported to have
several allelic forms

FUNCTION
#description a cofactor in the inactivation of C3b by serine proteinase I;
also increases the rate of dissociation of the C3bBb
complex (C3 convertase) and the (C3b)nb complex (C5
convertase) in the alternative complement pathway
#pathway complement alternate pathway
CLASSIFICATION
#superfamily complement factor H; complement factor H repeat
homology
KEYWORDS
#alternatively spliced; complement alternate pathway;
glycoprotein; plasma
FEATURE
1-18 #domain signal sequence #status predicted #label STG\
19-449 #product complement factor H, short splice form #status
experimental #label MAT\
21-80 #domain complement factor H repeat homology #label FH01\

```

	85-141	#domain complement factor H repeat homology #label FH0
	146-205	#domain complement factor H repeat homology #label FH0
	210-262	#domain complement factor H repeat homology #label FH0
	246-248	#region cell attachment (R-G-D) motif\
	267-320	#domain complement factor H repeat homology #label FH0
	325-385	#domain complement factor H repeat homology #label FH0
	389-482	#domain complement factor H repeat homology #label FH0
	21-66,52-80,85-129,	
	114-141,146-192,	
	178-205,210-251,	
	237-262,267-309,	
	294-320,325-374,	
	357-385,389-431,	
	416-442	
	217	
SUMMARY	#length 449 #molecular-weight 51007 #checksum 6077	
Query Match	100.0%; Score 2539; DB 1; Length 449;	
Best Local Similarity	100.0%; Pred. No. 0.00e+00;	
Matches 329;	Conservative 0; Mismatches 0; Indels 0; Gaps 0	
Dn	19 EDCNELPPRRNEILLGSSWDQTYREBGQAIIKCKPGRISLGNVIMCRKGEMWALNPLR 78	
Qy	1 EDENELPPRRNEILLGSSWDQTYREBGQAIIKCKPGRISLGNVIMCRKGEMWALNPLR 60	
Dn	79 KCQRPCGHGPDTPGTFTLLTGANFEYGVAVYCNNGYOLLLEINVRCDTDGMNDI 138	
Qy	61 KCQRPCGHGPDTPGTFTLLTGANNFEEIGVAVYTICNGYOLLLEINVRCDTDGMNDI 120	
Dn	139 PICEVVKCLPYTAPENGKIVSSAMEPDREYHFGAAVFVCNSGYIKIGDEMHCSDDGFV 198	
Qy	121 PICEVVKCLPYTAPENGKIVSSAMPDREYHFGAAVFVCNSGYIKIGDEMHCSDDGFV 180	
Dn	199 SKEKPACVEISCSPDYINGSPISOKITIKKENEFQYCKNMNGYSYSRGAVCTESGWRP 258	
Qy	181 SKEKPACVEISCSPDYINGSPISOKITIKKENRFQYCKNMNGEYSRGAVCTESGWRP 240	
Dn	259 LPSCEKSCDNPIYINGDISPLRIKHRTDEITYQCNRNGEYPATRGNTACTSTGWIPAP 318	
Qy	241 LPSCEKSCDNPIYINGDISPLRIKHRTDEITYQCNRNGEYPATRGNTACTSTGWIPAP 300	
Dn	319 RCTLKCDYPDRIKHGGLYHENMRPRPEPV 347	
Qy	301 RCTLKCDYPDRIKHGGLYHENMRPRPEPV 329	
RESULT	2	
ENTRY	NBHUH	#type complete
TITLE	#formal name Homo sapiens #common-name man	
ORGANISM	31-Dec-1993 #sequence-revision 31-Dec-1993 #text-change	
DATE	22-Jun-1999	
ACCSSIONS	S00254; A60238; A54726; A61565; A26505; I72654; S66298	
REFERENCE	S00254	
#authors	Ripoche, J.; Day, A.J.; Harris, T.J.R.; Slim, R.B.	
#journal	Biochem. J. (1988) 249:593-602	
#title	The complete amino acid sequence of human complement factor H.	
#cross-references	MOTID:88134059	
#accession	S00254	
#molecule-type	mRNA	
##residues	1-1231 ##label RIP	
##cross-references	EMBL:Y00716; NID:g931964; PIDN:CMA68704.1; PID:g931965	
#note	402-tyr was also found	
#note	parts of this sequence, including the amino and carboxyl ends of the mature protein, were confirmed by protein sequencing	
REFERENCE	A60238	Estallier, C.; Schwaeble, W.; Dierich, M.; Weiss, F.H.
#authors	Eur. J. Immunol. (1991) 21:799-802	
#journal	Human complement factor H: two factor H proteins are derived from alternatively spliced transcripts.	

##cross-references MUID:91184292
#accession A60238
##status not compared with conceptual translation
##molecule-type mRNA
##residues 1-56:1177-1231 ##label EST
#note only portions of this 4.3 kilobase mRNA were sequenced

REFERENCE
#authors A54726
#journal Day, A.J.; Ripchoe, J.; Lyons, A.; McIntosh, B.; Harris, T.J.R.; Sim, R.B.
#title Biosci. Rep. (1987) 7:201-207
#cross-references GB:M17517; NID:9180497; PIDN:AAA52016.1; PID:9180498
#accession A54726
#status not compared with conceptual translation
##molecule-type mRNA
##residues 'DFRN', 579-1231 ##label DAY
#cross-references GB:M17517; NID:9180497; PIDN:AAA52016.1; PID:9180498
#note parts of this sequence were determined by protein sequencing

REFERENCE
#authors A61565
#journal Ripchoe, J.; Day, A.J.; Willis, A.C.; Belt, K.T.; Campbell, R.D.; Sim, R.B.
#title Biosci. Rep. (1986) 6:65-72
#cross-references GB:M17517; NID:9180497; PIDN:AAA52016.1; PID:9180498
#accession A61565
#status not compared with conceptual translation
##molecule-type mRNA
##residues 'METGRNHLNAKI', 1050-1057, 'T', 1059-1102 ##label R12

REFERENCE
#authors A26505
#journal Sim, R.B.; Discolo, R.G.
#title Biochem. J. (1982) 205:285-293
#cross-references MUID:83048213
#accession A26505
#molecule-type protein
##residues 19-20, 'Q', 22-29, 'V', 31-33, 'Q', 35 ##label SIM

REFERENCE
#authors A4451
#journal Barlow, P.N.; Norman, D.G.; Steinkasserer, A.; Horne, T.J.; Pearce, J.; Driscoll, P.C.; Sim, R.B.; Campbell, I.D.
#title Biochemistry (1992) 31:3626-3634
#cross-references MUID:92232649
#accession A49224
#status annotation; NMR structure determination, residues 264-292

REFERENCE
#authors Norman, D.G.; Barlow, P.N.; Baron, M.; Day, A.J.; Sim, R.B.; Campbell, I.D.
#journal J. Mol. Biol. (1991) 219:717-725
#title Three-dimensional structure of a complement control protein module in solution.
#cross-references MUID:91278097
#accession I56100
#status annotation; NMR structure determination, residues 927-985

REFERENCE
#authors Esteller, C.; Koistinen, V.; Schwaible, W.; Dietrich, M.P.; Weiss, E.H.
#journal J. Immunol. (1991) 146:3190-3196
#title Cloning of the 1.4-kb mRNA species of human complement factor H reveals a novel member of the short consensus repeat family related to the carboxy terminal of the classical 150-kDa molecule.
#cross-references MUID:91201892
#accession I72654
#status translated from GB/EMBL/DBJ

REFERENCE
#molecule-type mRNA
##residues 1047-1231 ##label RES
#cross-references GB:M5294; NID:9183766; PIDN:AAA35948.1; PID:9183767
#accession S66298
#authors Caron, J.A.; Bates, R.C.; Smith, A.I.; Tetoz, T.; Arellano,

#journal A.; Gordon, D.L.; Burns, G.F.
#title Biochim. Biophys. Acta (1996) 1289:305-311
#cross-references MUID:96205365
#accession S66298
##status preliminary
##molecule-type protein
##residues 411-419:574-578, 580-582 ##label CAR
#note Factor H has also been found bound to cell membranes in an unknown manner. However, it has at least one cell attachment site motif in repeat 4.
#cross-references MUID:91201892
#accession I72654
#status translated from GB/EMBL/DBJ

REFERENCE
#authors Esteller, C.; Koistinen, V.; Schwaible, W.; Dietrich, M.P.; Weiss, E.H.
#journal J. Immunol. (1991) 146:3190-3196
#title Cloning of the 1.4-kb mRNA species of human complement factor H reveals a novel member of the short consensus repeat family related to the carboxy terminal of the classical 150-kDa molecule.
#cross-references MUID:91201892
#accession I72654
#status translated from GB/EMBL/DBJ

REFERENCE
#molecule-type mRNA
##residues 1047-1231 ##label RES
#cross-references GB:M5294; NID:9183766; PIDN:AAA35948.1; PID:9183767
#accession S66298
#authors Caron, J.A.; Bates, R.C.; Smith, A.I.; Tetoz, T.; Arellano,

[illegible]

```

#cross-references MUID:90111033
#accession 149711
#status preliminary; translated from GB/EMBL/DBD
#molecule_type mRNA
#residues 1-18 ##label RES
#cross-references GB:M31979; NID:g193726; PIDN:AAA37762.1; PID:g193729
REFERENCE 149728
#authors Munoz-Canoves, P.; Tack, B.F.; Vlk, D.P.
#journal Biochemistry (1989) 28:9891-9897
#title Analysis of complement factor H mRNA expression:
Dexamethasone and IFN-gamma increase the level of H in L
cells.
#cross-references MUID:90148935
#accession 149728
#status preliminary; translated from GB/EMBL/DBD
#molecule_type mRNA
#residues 1-19 ##label RE2
#cross-references GB:J2891; NID:g193805; PIDN:AAA37795.1; PID:g553926
COMMENT Two codonant alleles of factor H are present in mice.
COMMENT Factor H functions as a cofactor in the inactivation of C3b by
serine proteinase I and also increases the rate of dissociation
of the C3bb complex (C3 convertase) and the (C3b)nbb complex (C3
convertase) in the alternative complement pathway.
GENETICS
#map position 1
CLASSIFICATION #superfamily complement factor H; complement factor H repeat
homology
KEYWORDS complement alternate pathway; duplication; glycoprotein;
plasma
FEATURE
1-18 #domain signal sequence #status predicted #label SIG\
19-1234 #product complement factor H #status predicted #label
MPT\
21-80 #domain complement factor H repeat homology #label FH01\
85-141 #domain complement factor H repeat homology #label FH02\
146-205 #domain complement factor H repeat homology #label FH03\
210-262 #domain complement factor H repeat homology #label FH04\
246-248 #region cell attachment (R-G-D) motif\
267-320 #domain complement factor H repeat homology #label FH05\
325-385 #domain complement factor H repeat homology #label FH06\
389-404 #domain complement factor H repeat homology #label FH07\
448-505 #domain complement factor H repeat homology #label FH08\
509-564 #domain complement factor H repeat homology #label FH09\
569-622 #domain complement factor H repeat homology #label FH10\
629-683 #domain complement factor H repeat homology #label FH11\
690-743 #domain complement factor H repeat homology #label FH12\
752-802 #domain complement factor H repeat homology #label FH13\
808-861 #domain complement factor H repeat homology #label FH14\
867-931 #domain complement factor H repeat homology #label FH15\
936-989 #domain complement factor H repeat homology #label FH16\
994-1048 #domain complement factor H repeat homology #label FH17\
1053-1107 #domain complement factor H repeat homology #label FH18\
1114-1168 #domain complement factor H repeat homology #label FH19\
1172-1233 #domain complement factor H repeat homology #label FH20\
21-66,52-80,85-129,
114-141,146-192,
178-205,210-251,
237-262,267-309,
294-320,325-374,
357-385,389-431,
416-442,448-494,
477-505,509-553,
536-564,569-610,
597-622,629-672,
658-683,690-732,
718-743,752-791,
780-802,808-850,
836-861,867-920,
906-931,936-978,
964-989,994-1037,
1023-1048,
1053-1096,
1082-1107,

```



```

TITLE      completion factor H-related protein 3k4/5G4 - mouse
            (fragment)
ORGANISM   #formal_name Mus musculus #common_name house mouse
DATE       27-Jul-1990 #sequence_revision 31-Dec-1993 #text_change
           20-Aug-1999
ACCESSIONS A33068; B35068; C35068; D35068; E35068; F35068; G35068
REFERENCE   A33070
#authors    Vlk, D.P.; Munoz-Canoves, P.; Kozono, H.; Martin, L.G.; Tack
            B.F.; Chaplin, D.D.
#journal     J. Biol. Chem. (1990) 265:3193-3201
#title       Identification and sequence analysis of four complement
            factor H-related transcripts in mouse liver.
#cross-references MUID:90153969
#accession   A35068
            #molecule_type mRNA
            ##residues          1-452 ##label VIK
#cross-references GB:M29010; NID:g192561; PIDN:AA37415.1; PID:g387128
            #note               Translation of the nucleotide sequence is not complete
CLASSIFICATION #superfamily complement factor H; complement factor H repeat
            homology
FEATURE
  28-81      #domain complement factor H repeat homology #label FH1
  86-146     #domain complement factor H repeat homology #label FH2
  150-203    #domain complement factor H repeat homology #label FH3
  209-263    #domain complement factor H repeat homology #label FH4
  270-325    #domain complement factor H repeat homology #label FH5
  332-386    #domain complement factor H repeat homology #label FH6
  390-451    #domain complement factor H repeat homology #label FH7
SUMMARY     #length 452 #checksum 2383

Query Match 17.3%; Score 438; DB 2; Length 452;
Best Local Similarity 56.4%; Pred. No. 8,46e-78;
Matches 53; Conservativity 16; Mismatches 25; Indels 0; Gaps 0;

Db 15 TAMLTSTANGEEKTSCPPYLINGIYTPRIHHSRDSERYCNGEFPVNGSVSKCTPG 74
Qy 236 SSMRPRLPSCSEKSCNPPIPNMGDPSPLRIKHRTIDEITTCRNMGFPAIRGNTRAKTSTG 295
||| ||||||||:::|||||:::|||||
Db 75 WIPVRCTLKPCEEPEQEKYGRLYEESLRPFNY 108
Qy 296 WIPARCTLKPCDDYPDIKHGLYHENNRRPYFPY 329
||| ||||||||:::|||||:::|||||

RESULT      8
ENTRY        H35068 #type complete
TITLE        apolipoprotein H-related protein 23kL1 - mouse
ORGANISM     #formal_name Mus musculus #common_name mouse
DATE         27-Jul-1990 #sequence_revision 31-Dec-1993 #text_change
           10-Sep-1997
ACCESSIONS   A35068; A35069; B35069; I35068
REFERENCE     A35070
#authors      Vlk, D.P.; Munoz-Canoves, P.; Kozono, H.; Martin, L.G.; Tack
              B.F.; Chaplin, D.D.
#journal      J. Biol. Chem. (1990) 265:3193-3201
#title        Identification and sequence analysis of four complement
              factor H-related transcripts in mouse liver.
#cross-references MUID:90153969
#accession    H35068
              ##status          Preliminary
              ##molecule_type mRNA
              ##residues          1-303 ##label VIK
#cross-references GB:M29007; NID:g192557; PID:g309164; GB:J05259
              #note             translation of the nucleotide sequence is not complete
CLASSIFICATION #superfamily complement factor H repeat homology
            duplication
FEATURE
  28-81      #domain complement factor H repeat homology #label FH01
  86-146     #domain complement factor H repeat homology #label FH02
  150-203    #domain complement factor H repeat homology #label FH03
  212-266    #domain complement factor H repeat homology #label FH04
SUMMARY     #length 303 #molecular_weight 34498 #checksum 2835

Query Match 17.1%; Score 433; DB 2; Length 303;

```


	GENETICS	GDB:C4BPA
#gene	#cross-references GDB:120566; OMIM:120830	
#map-position	1q32-1q32	
#introns	48/1; 110/1; 143/2; 172/1; 236/1; 297/1; 362/1; 425/1; 482/1 540/3	
COMPLEX CLASSIFICATION	octamer of seven alpha chains and one beta chain	
KEYWORDS	#superfamily C4b-binding protein alpha chain; complement factor H repeat homology	
	acute phase; chylomicron; complement pathway; duplication; glycoprotein; plasma	
FEATURE		
1-48	#domain signal sequence #status predicted #label SIG	
49-597	#product C4b-binding protein alpha chain #status predicted #label MATN	
50-108	#domain complement factor H repeat homology #label FH1	
113-170	#domain complement factor H repeat homology #label FH2	
175-234	#domain complement factor H repeat homology #label FH3	
239-294	#domain complement factor H repeat homology #label FH4	
299-360	#domain complement factor H repeat homology #label FH5	
364-422	#domain complement factor H repeat homology #label FH6	
381-404	#region complement C4b binding #status predicted	
426-480	#domain complement factor H repeat homology #label FH7	
484-538	#domain complement factor H repeat homology #label FH8	
221,506,528	#binding_site carbohydrate (Asn) (covalent) #status experimental	
SUMMARY	#length 597 #molecular-weight 67033 #checksum 6374	
Query Match	17.1%; Score 433; DB 1; Length 597;	
Best local Similarity	29.8%; Pred. No. 1,45e-76;	
Matches	97; conservative 68; Mismatches 132; Indels 29; Gaps 26;	
Db	67 TERREFTGTLTKTCLPGVYRSHSTQLTGNSDGEW-YNTFCIY-KR-CRHPGLRNQG 123	
Oy	20 SDQYPEGGQAIVKCRPGY-RSLGNVIMVCRK-GEMVALNPILRCQKRGPHDPDTFGT 77	
Db	124 VEIKT-DL-SFSGQIDPESCEGFLLGSTSRCEVDQRGVGMSHPLQCELVKCKPPDI 181	
Oy	78 FLITGNGVEYGKANYTCNEGTLGELIYNR-EC-DTD-GWTNDIPICEVVKCLPTAP 134	
Db	182 KNGR--HSG-E-ENFYAYGSVTYSCDPRFLSLGHASISCTVENETIGVWRPSPPTCEKI 237	
Oy	135 ENGIIVSAMEPDRREHFQGAVFVNCNGVKGIDDEMHMS-D-D-GFMSEKPKVEI 190	
Db	238 TORKPDVSGEVMVSGGRPIYNYKDITVFKCQKGFVLRGSSVIHCDAADSKWNPSPACEPN 297	
Oy	191 SKSPDVINGSPISO-KIIYEKENRFQYCNMGXEYSERDDAVC-TESGWRP-LPSCEEK 247	
Db	298 SCINLPDIPHASMETYPRTKEDEVYVAVLRFRCHRGVGYPTTDEPTTVICQNLKRPY 357	
Oy	248 SCDN-PYIPNG--D-YS-PLRIK-HHTGDIEITQCXNGGFPAPRGMTA-KC-TSTGWIPA 299	
Db	358 QCGCALCPPEKLNGGEITQHRRKSP 383	
Oy	300 PRCITLKPCDYDPDIKHGILYHENMRP 325	
RESULT 10		
ENTRY	I36936 #type fragment	
TITLE	Complement receptor 1 - chimpanzee (fragment)	
ORGANISM	#formal_name Pan troglodytes #common_name chimpanzee	
DATE	02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 13-Aug-1999	
ACCESSIONS	I36936	
REFERENCE	I36936	
#authors	Birmingham, D.J.; Shen, X.P.; Hourcade, D.; Nickells, M.W.; Atkinson, J.P.	
#journal	J. Immunol. (1994) 153:691-700	
#title	Primary sequence of an alternatively spliced form of CR1. Candidate for the 75,000 Mr complement receptor expressed on chimpanzee erythrocytes.	
#cross-references	MIM:94292799	
#accession	I36936	

##status	preliminary: translated from GB/EMBL/DBS
##molecule_type	mRNA
##cross-references	1-2014 ##label RES
CLASSIFICATION	GB:U24920; NID:9551564; PID:9557725
FEATURES	##cross-references
18-74	##suprafamily complement factor H repeat homology
79-136	#domain complement factor H repeat homology #label F1
141-207	#domain complement factor H repeat homology #label F1
213-268	#domain complement factor H repeat homology #label F1
273-328	#domain complement factor H repeat homology #label F1
333-391	#domain complement factor H repeat homology #label F1
396-462	#domain complement factor H repeat homology #label F1
470-536	#domain complement factor H repeat homology #label F1
529-586	#domain complement factor H repeat homology #label F1
722-778	#domain complement factor H repeat homology #label F1
1041-1107	#domain complement factor H repeat homology #label F1
1172-1228	#domain complement factor H repeat homology #label F1
1233-1291	#domain complement factor H repeat homology #label F1
1291-1362	#domain complement factor H repeat homology #label F1
1326-1382	#domain complement factor H repeat homology #label F1
1432-1489	#domain complement factor H repeat homology #label F1
1625-1681	#domain complement factor H repeat homology #label F1
1749-1815	#domain complement factor H repeat homology #label F1
1823-1879	#domain complement factor H repeat homology #label F1
1884-1940	#domain complement factor H repeat homology #label F1
SUMMARY	#length 2014 #checksum 672
Query Match	16.7%: Score 423; DB 2; Length 2014;
Best Local Similarity	29.1%: Pred. No.4.26e-74;
Matches	95; Conservative 74; Mismatches 119; Indels 39; Gaps 2
Db	1391 PFVGTSLNYEGRPEY--FGKMFISICLENLWSSYED--NCRKSCGPPE-PFNGVYHI 1445
OY	24 YPEGQATLYKCRPSRYSLGNVYV--CRKG-EMVALNPLKCKQKRCRGGHDPFOTETLT 81
Db	1446 NTDI--QFGSYVNSCNGEFLILISPTGLVGSNNYTWKAPLFEIISCEPPLISNGD 1504
OY	82 GGNVEYGVAVYTCNNGYDLDE--INTRECDTD-GWINDPILEVKCLPYAPENCK 138
Db	1505 FYSNNRAS--FHNGIVVYTCHTGPDGDLFEIVGERSIYCTSKDPOGVWSSPPPCI 1561
OY	139 IVSSAMPDEDEYHGGAVRFVNSG-----YKIGDDEMHG-S-DD--GWSKEXKRCV 188
Db	1562 STNCTAPVEYENAIKVPDGNSSPFSLEIYRFGCQGFYVAGVSHYVOCOTNGRWPKLPHC 1621
OY	189 EIS--CKSPVDINSSPIS--OKIITYKENERYQYKNNGYEYSEKGAIVCTESG--WRP-LPSC 244
Db	1622 S-RVCCPPELILGEHTPSHODNFSGQVFNFSCEPG-Y-DLRGAASLHCPQDMSPEA 1678
OY	245 EKSQCDNP-YTPRGDVSPLRIKH-RGDDITTYQCANGFYPATRGNTA-KCTSTS-WTP-A 299
Db	1679 PRCIVKSCDFFLGQLDPRHAYPLPNIQ 1705
OY	300 PRCITLKPCD-Y-PDIKHG-GLYHENMR 323
RESULT	11
ENTRY	A28507
TITLE	complement C3b/C4b receptor precursor, membrane-bound (allotype F) - human
ALTERNATE_NAMES	complement receptor type 1 (CRI); surface glycoprotein CD35
ORGANISM	Human Homo sapiens #common name man
DATE	19-Nov-1988 #sequence_revision 06-Sep-1996 #text-change 10-Sep-1997
ACCSIONS	S03843; A28507; I56203; A24748; B24748; C24748
REFERENCE	S03843
#authors	Klickstein, L.B.; Bartow, T.J.; Miletic, V.; Rabson, L.D.; Smith, J.A.; Fearon, D.T.
#journal	J. Exp. Med. (1988) 168:1699-1717
#title	Identification of distinct C3b and C4b recognition sites in the human C3b/C4b receptor (CRI, CD35) by deletion mutagenesis.
#cross-references	MUTID:89035992

	ENTRY	15	NBMSCA	#type complete
	TITLE		C4b-binding protein alpha chain precursor - mouse	
	ALTERNATE_NAMES		C4BP; proline-rich protein	
	ORGANISM		#formal_name Mus musculus #common_name house mouse	
	DATE		30-Jun-1993 #sequence_revision 30-Jun-1993 #text_Change 18-Jul-1997	
	ACCESSIONS		A27117	
	REFERENCE		A27117	
	#authors		Kristensen, T.; Ogata, R.T.; Chung, L.P.; Reid, K.B.M.; Tack, B.F.	
	#journal		Biochemistry (1987) 26:4668-4674	
	#title		cDNA structure of murine C4b-binding protein, a regulatory component of the serum complement system.	
	#cross-references		MUID:88024997	
	#accession		A27117	
	##residues		#molecule_type mRNA	
	##note		1-469 ##label KRI the authors translated the codon GGT for residue 25 as Val	
	COMMENT		C4BP controls the classical pathway of complement activation. It binds as a cofactor to C3b/C4b inactivator (C3bINA), which then hydrolyzes the complement fragment C4b. It also accelerates the degradation of the C4b2a complex (C3 convertase) by dissociating the complement fragment C2a. In mouse, C4BP is a multimeric protein of noncovalently associated chains.	
	CLASSIFICATION		#superfamily C4b-binding protein alpha chain; complement factor H repeat homology	
	KEYWORDS		acute phase; chylomicron; complement pathway; duplication; glycoprotein; plasma	
	FEATURE			
	1-56		#domain signal sequence #status predicted #label SIG\	
	57-469		#product C4b-binding protein alpha chain #status predicted #label MAT\	
	58-115		#domain complement factor H repeat homology #label FH1\	
	120-116		#domain complement factor H repeat homology #label FH2\	
	181-240		#domain complement factor H repeat homology #label FH3\	
	245-289		#domain complement factor H repeat homology #label FH4\	
	303-355		#domain complement factor H repeat homology #label FH5\	
	359-413		#domain complement factor H repeat homology #label FH6\	
	74,227,275,292,366,		#binding_site carbohydrate (Asn) (covalent) #status predicted	
	381,428		#length 469 #molecular_weight 51523 #checksum 5359	
SUMMARY				
Query Match	15.8%;	Score 402;	DB 1; Length 469;	
Best Local Similarity	27.9%;	Pred. No. 6,13e-69;		
Matches	92;	Conservative 74;	Mismatches 133; Indels 31; Gaps 28;	
Dd	60	PPPAIPNALPADVNRKTFESHHTLLTKYLCLELGGYRGISRMWYCKPSGEM-EIS-VS-CA	116	
		: : : : : : : : : : : : : : :		
Oy	7	PPRRNREILTGSMSDQT-YPECTQAIKYCRPGY-RSLGNAINMCCR-KGEVWVAIPLMKKO	63	
Dd	117	KKKCRNPGLINDG-Y-VNCERT-TTGQSIEESCOEGFLIVSSR--SCSEVRKGAVASNP	172	
		: :		
Oy	64	KRCGCHPAGTPTGFTFLVGNAVFEYGVKAAYITCNIGVILLSEINRYRCDI-D-G--WTND	119	
Dd	173	FPEPCVIYKGPDPDISNGK-HSGT-E-DF-YPYNHNSITYTCDDPGRFLVGSPIFGCTIVNK	228	
		: : : : : : : : : : : : : : : : : : : :		
Oy	120	IPICEYVKCLPYAPRNKGATIVSSAMEPRREHFQGAVFVCNSGYKTLEGDEMHCS--DD	177	
Dd	229	TVPWSSSPPTCEKIICOSPNIHLGVYISGYKATYTHDSVRLACLNGTLVLRGHVIEQC	288	
		: : : : : : : : : : : : : : : : : : :		
Oy	178	GF-WSEKERKCVEISKSPDIVINSPTSQ-KIIKENERFOYCCNMNGEYESENGDAVCT	234	
Dd	289	GNGWSSLPICEF-DDDLPPAIVNGYYSM-VYSKIT-LVTYECDKGYRLGAKIIS-CS	344	
		: : : : : : : : : : : : : : : : : : : :		
Oy	235	ESG-WRLPLPCEKESCD-NPIIPNGYISPLRKIKHTGTGETIYTCCNGFYPATRGNTAKCT	292	
Dd	345	FSKWKGTAPOCKA-LCOKPELVNGTILSDCK	373	

Fri Jun 9 10:53:40 2000

US-09-316-163-11.rpr

Page 12

QY 293 STGWI-PAPROCTLRKPCDYPDIKHGGLYHEN 321

Search completed: Thu Jun 8 21:45:47 2000
Job time : 23 secs.

 M P E R E H
 (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution Rights by Oxford Molecular Ltd

Mpsrch_p protein - protein database search, using Smith-Waterman algorithm
 Run on: Thu Jun 8 21:46:04 2000; Maspar time 12.45 Seconds
 Tabular output not generated. 804,485 Million cell updates/sec

Title: >US-09-316-163-11
 Description: (1-329) from US09316163.pep
 Perfect Score: 2539
 Sequence: 1 EDCNEUPPRNTEITLGMSW.....PIKKGGLGHENMRPRYPV 329

Scoring table: PAM 150
 Gap 11

Searched: 83857 seqs, 30454973 residues

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database: Swiss-Prot
 1:Swissprot

Statistics: Mean 44.764; Variance 62.937; scale 0.711

Pred. No. 18, the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	2539	100.0	1231	1	CEFAH_HUMAN COMPLEMENT FACTOR H PR	0.00e+00
2	1772	69.8	1234	1	CEFAH_MOUSE COMPLEMENT FACTOR H PR	0.00e+00
3	433	17.1	1597	1	C4BP_HUMAN C4B-BINDING PROTEIN AL	1.94e-86
4	409	16.1	2039	1	CRL_HUMAN COMPLEMENT RECEPTOR TY	8.70e-80
5	404	15.9	538	1	C4BP_RAT C4B-BINDING PROTEIN AL	2.09e-78
6	402	15.8	469	1	C4BP_MOUSE C4B-BINDING PROTEIN AL	7.44e-78
7	388	15.3	263	1	VCP_VACCV COMPLEMENT CONTROL PRO	5.28e-74
8	376	14.8	830	1	LEM3_HUMAN P-SELECTIN PRECURSOR (1.02e-70
9	371	14.6	345	1	APOH_MOUSE BETA-2-GLYCOPROTEIN I	2.38e-69
10	367	14.5	768	1	LEM3_RAT P-SELECTIN PRECURSOR (2.94e-68
11	363	14.3	668	1	F13B_MOUSE COAGULATION FACTOR XII	3.61e-67
12	357	14.1	610	1	C4BP_BOVIN C4B-BINDING PROTEIN AL	1.54e-65
13	343	13.5	768	1	LEM3_MOUSE P-SELECTIN PRECURSOR (9.54e-62
14	341	13.4	360	1	CCPH_HYVSA COMPLEMENT CONTROL PRO	3.31e-61
15	340	13.4	661	1	F13B_HUMAN COAGULATION FACTOR XII	6.15e-61
16	334	13.2	345	1	APOH_BOVIN BETA-2-GLYCOPROTEIN I	2.54e-59
17	333	13.2	377	1	MCP_HUMAN MEMBRANE COPOLYMER PROT	4.72e-59
18	330	13.1	345	1	APOH_HUMAN BETA-2-GLYCOPROTEIN I	4.72e-59
19	320	12.6	624	1	LEM2_HUMAN E-SELECTIN PRECURSOR (1.44e-55
20	314	12.4	612	1	LEM2_MOUSE E-SELECTIN PRECURSOR (7.75e-54
21	310	12.2	958	1	HIG_DROME LOCOMOTION-RELATED PRO	6.68e-53
22	306	12.1	345	1	APOH_CANFA BETA-2-GLYCOPROTEIN I	7.71e-52
23	308	12.1	484	1	LEM2_PIG E-SELECTIN PRECURSOR (2.27e-52

24	306	12.1	551	1	LEM2_RABIT E-SELECTIN PRECURSOR (7.71e-52
25	306	12.1	769	1	LEM3_SHEEP P-SELECTIN PRECURSOR (7.71e-52
26	302	11.9	1033	1	CR2_HUMAN COMPLEMENT RECEPTOR TY	8.86e-51
27	299	11.8	330	1	CEPH_HUMAN COMPLEMENT FACTOR H-LI	5.51e-50
28	299	11.8	549	1	LEM2_RAT E-SELECTIN PRECURSOR (5.51e-50
29	299	11.8	611	1	LEM2_CANFA E-SELECTIN PRECURSOR (5.51e-50
30	293	11.5	381	1	DAF_HUMAN COMPLEMENT DECAV-ACCEL	2.11e-48
31	289	11.4	1025	1	CR2_MOUSE COMPLEMENT RECEPTOR TY	2.38e-47
32	284	11.2	340	1	DAF_PONPY COMPLEMENT DECAV-ACCEL	4.89e-46
33	281	11.1	331	1	CFPH_HUMAN COMPLEMENT FACTOR H-LI	2.98e-45
34	280	11.0	507	1	DAF_CAVPO COMPLEMENT DECAV-ACCEL	5.44e-45
35	276	10.9	287	1	APOH_RAT BETA-2-GLYCOPROTEIN I	6.03e-44
36	276	10.9	495	1	LEM2_BOVIN E-SELECTIN PRECURSOR (6.03e-44
37	276	10.9	1019	1	LFC_TACTR LIMULUS CLOTTING FACTO	6.03e-44
38	271	10.7	646	1	LEM3_BOVIN P-SELECTIN PRECURSOR (1.91e-42
39	254	10.0	390	1	DAF1_MOUSE COMPLEMENT DECAV-ACCEL	2.95e-38
40	238	9.4	407	1	DAF2_MOUSE COMPLEMENT DECAV-ACCEL	3.52e-34
41	232	9.1	372	1	LEM1_RAT L-SELECTIN PRECURSOR (1.15e-32
42	231	9.1	372	1	LEM1_MOUSE L-SELECTIN PRECURSOR (1.15e-32
43	220	8.7	372	1	LEM1_PANTR L-SELECTIN PRECURSOR (1.15e-29
44	220	8.7	372	1	LEM1_MACMU L-SELECTIN PRECURSOR (1.15e-29
45	220	8.7	372	1	LEM1_HUMAN L-SELECTIN PRECURSOR (1.15e-29

ALIGNMENTS

RESULT ID	1	CEFAH_HUMAN	STANDARD;	PRT;	1231 AA.
AC	P08603;				
DT	01-AUG-1988 (Rel. 08, Created)				
DT	01-JUN-1990 (Rel. 13, Last sequence update)				
DT	15-FEB-2000 (Rel. 39, Last annotation update)				
DE	COMPLEMENT FACTOR H PRECURSOR.				
GN	HPI OR HP OR CFH.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
RN	[1]	SEQUENCE FROM N.A.			
RP	TISSUE-LIVER:				
RC	MEDLINE: 88134059.				
RX	Rapoche J., Day A.J., Harris T.J.R., Sim R.B.;				
RA	"The complete amino acid sequence of human complement factor H.;"				
RT	Biochem. J. 249:559-602(1988).				
RL	[2]	SEQUENCE OF 53-445 FROM N.A.			
RN	MEDLINE: 87054207.				
RP	Schulz T.F., Schwaeble W., Stanley K.K., Weiss E., Dierich M.P.;				
RA	"Human complement factor H: isolation of cDNA clones and partial cDNA sequence of the 38-kDa tryptic fragment containing the binding site for C3b.;"				
RT	Eur. J. Immunol. 16:1351-1355(1986).				
RL	[3]	SEQUENCE OF 226-445 FROM N.A., AND PARTIAL SEQUENCE.			
RN	MEDLINE: 86169701.				
RX	Kristensen T., Wetzel R.A., Tack B.F.;				
RA	"Structural analysis of human complement protein H: homology with C4b binding protein, beta 2-glycoprotein I, and the Ba fragment of B2.;"				
RT	J. Immunol. 136:3407-3411(1986).				
RL	[4]	SEQUENCE OF 1047-1231 FROM N.A.			
RN	MEDLINE: 91201892.				
RP	Estallier C., Koistinen V., Schwaeble W., Dierich M.P., Weiss E.H.;				
RA	"Cloning of the 1.4-kb mRNA species of human complement factor H reveals a novel member of the short consensus repeat family related to the carboxy terminal of the classical 150-kDa molecule.;"				
RT	J. Immunol. 146:3190-3196(1991).				
RL	[5]	SEQUENCE OF 19-35.			
RN	MEDLINE: 83048213.				
RP	Sim R.B., Discipio R.G.;				
RA	"Purification and structural studies on the complement-system control protein beta 1H (Factor H).;"				
RT					

DR	EMBL; X04284;	CAB51244.1;	-
DR	EMBL; X04296;	CAA27839.1;	-
DR	EMBL; X02865;	CA26617.1;	-
DR	PIR; A33568;	NBHUC4.	
DR	HSSP; P10998;	IYVC.	
DR	MIM; 120830;	-	
DR	PFAM; PF00084;	sushi; 8.	
KM	Complement pathway; Plasma; Glycoprotein; Repeat; Sushi; Signal; Polymorphism.		
FT	SIGNAL	1	48
FT	CHAIN	49	597
FT	DOMAIN	49	539
FT	REPEAT	49	109
FT	REPEAT	112	171
FT	REPEAT	174	235
FT	REPEAT	238	295
FT	REPEAT	298	361
FT	REPEAT	364	423
FT	REPEAT	425	481
FT	REPEAT	483	539
FT	DISULFID	50	96
FT	DISULFID	81	108
FT	DISULFID	113	154
FT	DISULFID	140	170
FT	DISULFID	175	217
FT	DISULFID	203	234
FT	DISULFID	269	281
FT	DISULFID	297	294
FT	DISULFID	329	348
FT	DISULFID	332	360
FT	DISULFID	364	2387
FT	DISULFID	7399	7409
FT	DISULFID	426	422
FT	DISULFID	454	468
FT	DISULFID	484	525
FT	DISULFID	511	538
FT	DISULEID	546	546
FT	DISULEID	538	558
FT	CARBOHYD	221	221
FT	CARBOHYD	506	506
FT	CARBOHYD	528	528
FT	VARIANT	92	92
FT	O -> T. FTID=VAR_001977.		
FT	VARIANT	357	357
FT	Y -> H. FTID=VAR_001978.		
FT	VARIANT	357	357
FT	/FTID=VAR_001978.		
SO	SEQUENCE	597 AA;	67033 MW; 67E03FE2A85A16DD CRC64;
<hr/>			
Query Match 17.1%; Score 433; DB 1; Length 597;			
Best Local Similarity 29.8%; Pred. No. 1,94e-86;			
Matches 97; Conservative 68; Mismatches 132; Indels 29; Gaps 26;			
<hr/>			
Db	67 TETPRKGTGLIKYTCLPGVVRSHSTQTLTICNSDEMW-VNTECIY-KR-CRHGELRNQ	123	
Oy	20 SDQATPECTQAAYCRPSY-RSLGNVIWCKR-GEWALNLRCOKRPCGHDPDFGT	77	
Db	124 VEIK-DL-SHGSDIEFCSEGFLLIGSTRSCVODRGVMSHPLQCCEIVKKPPDI	181	
Oy	78 FTLLGGNVFEYGVAVIYCNGEIQLGEINRK-EC-DTD-GWTINDPICEVYKCLPVTA	134	
Db	182 RNGR--HSQG-E-ENFYAGFSVTYSCDPFRSILHASICTVENETIGVMRSPPTCEKI	237	
Oy	135 ENGRIVSAMPDREYHRHQVAFVFCVNGGYKIEBDEMHC-D-D--GFMEKEKPCVEI	190	
Db	238 TCRRKDVSHGHEVSGFGIYNKDTIVKCOKGEGVLKSSSVIHCDADSKMNPSPACEPN	297	
Oy	191 SCRSVDVINGSPIISO-KLIYEKENEFQYCKMGGEYSERGAIV-TESGWRP-LPSCEEK	247	
Db	298 SCINLPDIPHAMSWETYPRTKEDYVVGVTVLRCHGPYKPTTBPTVICOKMLMTPT	357	
Oy	248 SCDN-PYLPG--D-Y-S-PLRIK-HRTGDELTYCCRNFGYFATIGNTA-KC-ISTGMIPA	299	
Db	358 OGCEALCPPEPKLNGETITQRHKSHP	383	

```

0y      300 PROTLKPCDYPDIKHGFLHYENMRAP 325

RESULT      4
ID          CRI_HUMAN      STANDARD:      PRT:      2039 AA.
AC          P19727;
DT          01-NOV-1990 (Rel. 16, Created)
DT          01-NOV-1990 (Rel. 16, Last sequence update)
DT          01-NOV-1997 (Rel. 35, Last annotation update)
DE          COMPLEMENT RECEPTOR TYPE 1 PRECURSOR (C3B/C4B RECEPTOR) (CD35
DE          ANTIGEN).
OS          CRI OR C3BR.
OS          Homo sapiens (Human).
OC          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC          Eutheria; Primates; Catarrhini; Hominoideae; Homo.
RN          [1]
RX          SEQUENCE FROM N.A.
RX          MEDLINE; 89035992.
RA          Klickstein L.B., Bartow T.J., Miletic V., Rabson L.D., Smith J.A.,
RA          Fearon D.T.;
RT          "Identification of distinct C3b and C4b recognition sites in the
RT          human C3b/C4b receptor (CRI, CD35) by deletion mutagenesis.";
RL          J. Exp. Med. 168:1699-1717(1988).
RN          [2]
RN          SEQUENCE OF 503-2039 FROM N.A.
RX          MEDLINE; 87168191.
RA          Klickstein L.B., Wong W.W., Smith J.A., Weis J.H., Wilson J.G.,
RA          Fearon D.T.;
RT          "Human C3b/C4b receptor (CRI). Demonstration of long homologous
RT          repeating domains that are composed of the short consensus repeats
RT          characteristics of C3/C4 binding proteins.";
RL          J. Exp. Med. 165:1095-1112(1987).
RN          [3]
RN          SEQUENCE OF 761-783; 831-845 AND 1179-1195 FROM N.A.
RX          MEDLINE; 86067975.
RA          Wong W.W., Klickstein L.B., Smith J.A., Weis J.H., Fearon D.T.;
RT          "Identification of a partial cDNA clone for the human receptor for
RT          complement fragments C3b/C4b.";
RL          Proc. Natl. Acad. Sci. U.S.A. 82:7711-7715(1985).
CC          -I- FUNCTION: CRI, PRESENT ON ERYTHROCYTES, LEUKOCYTES, GLOMERULAR
CC          PODOCYTES, AND SPLENIC FOLLICULAR DENDRITIC CELLS. MEDIATES THE
CC          BINDING BY THESE CELLS OF PARTICLES AND IMMUNE COMPLEXES THAT HAVE
CC          ACTIVATED COMPLEMENT.
CC          *
CC          -I- SUBUNIT: MONOMER.
CC          -I- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC          -I- POLYMORPHISM: CRI IS RESPONSIBLE FOR THE KNOPS BLOOD GROUP SYSTEM.
CC          -I- MISCELLANEOUS: SEVEN SHORT CONSENSUS REPEATS (SCR) CONSTITUTE A
CC          LONG HOMOLOGOUS REPEAT (LHR). THE N-TERMINAL TWO SRCS OF LHR-A
CC          CONTAINED A SITE DETERMINING C4 SPECIFICITY, AND THE N-TERMINAL
CC          TWO SRCS OF LHR-B AND -C EACH HAD A SITE DETERMINING C3
CC          SPECIFICITY.
CC          -I- MISCELLANEOUS: THIS IS THE SEQUENCE OF THE F ALLOTYPIC CRI.
CC          -I- SIMILARITY: CONTAINS 30 SUSHI (SCR) REPEATS.
CC          -----
CC          THIS SWISS-PROT entry is copyright. It is produced through a collaboration on
CC          between the Swiss Institute of Bioinformatics and the EMBL outstation
CC          the European Bioinformatics Institute. There are no restrictions on
CC          use by non-profit institutions as long as its content is in no way
CC          modified and this statement is not removed. Usage by and for commercial
CC          entities requires a license agreement (see http://www.isb-sib.ch/announcement/).
CC          or send an email to license@isb-sib.ch).
CC          -----
DR          EMBL; M11569; AAA52297.1; -
DR          EMBL; M11617; AAA52298.1; -
DR          EMBL; M11618; AAA52299.1; -
DR          EMBL; Y00816; CAA68755.1; -
DR          EMBL; X05309; CAA28933.1; -
DR          PIR; A28507; A28507.
DR          PIR; A24748; A24748
DR          PIR; B24748; B24748
DR          PIR; C24748; C24748
DR          PIR; S03843; S03843.

```


DR	PFAM; PF00008; EGF. 1.	
DR	PFAM; PF00059; lectin. C. 1.	
KW	Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin	
KW	Selectin; Signal; Sushi; Repeat; Polymorphism; 3d-structure.	
FT	SIGNAL	1
FT	CHAIN	42
FT	DOMAIN	42
FT	TRANSMEM	772
FT	DOMAIN	796
FT	DOMAIN	58
FT	DOMAIN	159
FT	DOMAIN	199
FT	REPEAT	199
FT	REPEAT	261
FT	REPEAT	323
FT	REPEAT	385
FT	REPEAT	447
FT	REPEAT	509
FT	REPEAT	571
FT	REPEAT	641
FT	REPEAT	703
FT	DISULFID	60
FT	DISULFID	131
FT	DISULFID	163
FT	EXTRACELLULAR (POTENTIAL).	830
FT	POTENTIAL.	771
FT	CYTOPLASMIC (POTENTIAL).	795
FT	C-TYPE LECTIN (SHORT FORM).	830
FT	EGF-LIKE.	158
FT	9 X SUSHI (SCR) REPEATS.	195
FT	SUSHI 1.	762
FT	SUSHI 2.	258
FT	SUSHI 3.	320
FT	SUSHI 4.	382
FT	SUSHI 5.	444
FT	SUSHI 6.	506
FT	SUSHI 7.	568
FT	SUSHI 8.	630
FT	SUSHI 9.	700
FT	BY SIMILARITY.	762
FT	BY SIMILARITY.	158
FT	BY SIMILARITY.	174
FT	BY SIMILARITY.	174

[illegible]

ID	AP0H_MOUSE	STANDARD:	PRT:	345 AA.
AC	001339:			
DT	01-APR-1993 (rel. 25, Created)			
DT	01-APR-1993 (rel. 25, Last sequence update)			
DE	15-JUL-1999 (rel. 38, Last annotation update)			
DE	BETA-2-GLYCOPROTEIN I PRECURSOR (APOLIPOPROTEIN H) (APO-H) (ACTIVATED			
DE	PROTEIN C-BINDING PROTEIN) (APC INHIBITOR) (B2GPI).			
GN	AP0H OR B2GPI.			
OS	Mus musculus (Mouse)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;			
OC	Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Medline; 92372000.			
RA	Nonaka M., Matsuda Y., Shirosaki T., Moriwa K., Natsune-Sakai S.;			
RT	"Molecular cloning of mouse beta 2-glycoprotein I and mapping of the			
RT	gene to chromosome 11."			
RL	Genomics 13:1082-1087(1992).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-CBA/J; TISSUE-LIVER;			
RX	Medline; 94242017.			
RA	Seller G.C., Steel D.M., Zafiroopoulos A., Seery L.T.,			
RA	Whitehead A.S.;			
RT	"Characterization, expression and evolution of mouse beta 2-			
RT	glycoprotein I (apolipoprotein H)."			
RL	Biochem. Biophys. Res. Commun. 200:1521-1528(1994).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-BALB/C; TISSUE-LIVER;			
RA	Kristensen T.;			
RT	"Structure of the human beta-2-glycoprotein I gene.;"			
RL	Submitted (FEB-1997) to the EMBL/Genbank/DBJ databases.			
CC	SUCH AS HEPARIN, PHOSPHOLIPIDS, AND DEXTRAN SULFATE. MAY PREVENT			
CC	ACTIVATION OF THE INTRINSIC BLOOD COAGULATION CASCADE BY BINDING			
CC	TO PHOSPHOLIPIDS ON THE SURFACE OF DAMAGED CELLS.			
CC	-I- TISSUE SPECIFICITY: PLASMA.			
CC	-I- SIMILARITY: CONTAINS 4 SUSHI (SCR) REPEATS.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; D10056; BAA00945.1; -			
DR	EMBL; S70439; AAB30789.1; -			
DR	EMBL; Y11356; CAA72190.1; -			

[illegible]

FT	DISULFID	549	549	INTERCHAIN (WITH BETA CHAIN) (POTENTIAL).
FT	DISULFID	561	561	INTERCHAIN (WITH BETA CHAIN) (POTENTIAL).
FT	CARBOHYD	66	66	POTENTIAL.
FT	CARBOHYD	221	221	POTENTIAL.
FT	CARBOHYD	525	525	POTENTIAL.
FT	CARBOHYD	602	602	POTENTIAL.
SO	SEQUENCE	610 AA:	68886 MW:	D806B270E8A05B58 CRC64:
Query Match				
Best Local Similarity		14.1%;	Score 357;	DB 1: Length 610:
Matches		85; Conservative	Pred. No. 1.54e-65;	Indels 30; Gaps 25
Matches 85; Conservative				
Db	NETREFETGLIARTCPGGRYISRRKNFLICDGTDMW-KYKEF--CYVKRCENPGELNGQ	122		
Qy	SDQIYPGNGTALIKKCPGRY-SLGNVIWYCR-KGEWALNPLKCKQKRRCGHGDPNGET	77		
Db	123 VIYKT-D-YSGSEIFSCSESYVIGSANSYCOLQKGYWSDPLPQCLIKCEPPTI	180		
Qy	78 FTLLGGWVEFYGYKAVAYTCNEGYTLGGEIN-YRCDDDG--WINDPIDEYVKCLPTAP	134		
Db	181 SNRG-HNGG-DEEF-YTGSSTVYSCDDEFMCLKASISCRVENKIKGWSSPSCKKV	237		
Qy	135 ENGIYSSANEPREHFGQAFVFCNSGYIIEDEEMHS-DD---GFMSKKEPKCVBI	190		
Db	238 ICQVYVVKDKITSGFGPIYTYQOSIVYACKNGFRLBGDSLICEADNSWNPPTCEIN	297		
Qy	191 SKSPDYVINSPISQ-KIITYKENERFOYKCMNGEYSRGDAVC-TESGWRPLP-SCERK	247		
Db	298 GCLGPHIPIALMERDHOQTEQOVYDIGVLSYKCHFGFKPEPTDPTVTCQSNLEWS	357		
Qy	248 SC-DNPIIPMG--D-YG-PLRKHKRGD-E-I-TYOCRNGFYPTATRGNTA-KCTSI-GMI	297		
Db	358 FYIECKEYCCPEPNLN 374			
Qy	298 PAPRCTLKPDYDPDIKH 314			
RESULT 13				
ID	LEM3 MOUSE	STANDARD:	PRT:	766 AA.
AC	Q01102;			
DT	01-APR-1993 (Rel. 25, Created)			
DT	01-APR-1993 (Rel. 25, Last sequence update)			
DT	15-JUN-1999 (Rel. 38, Last annotation update)			
DE	P-SELECTIN PRECURSOR (GRANULE MEMBRANE PROTEIN 140) (GMP-140) (PADGEM			
DE	(C662P) (LEUCOCYTE-ENDOTHELIAL CELL ADHESION MOLECULE 3) (LECAM3).			
OS	SELP OR GMP.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;			
OC	Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE: 92340571.			
RA	Sanders W.E. Jr., Wilson R.W., Ballantyne C.M., Beaudet A.L.;			
RT	"Molecular cloning and analysis of in vivo expression of murine p-			
RT	selectin.";			
RL	J. Biol. Chem. 267:15176-15183(1992).			
CC	[2]			
CC	SEQUENCE FROM N.A.			
CC	MEDLINE: 92345617.			
CC	Sanders W.E. Jr., Wilson R.W., Ballantyne C.M., Beaudet A.L.;			
CC	"Molecular cloning and analysis of in vivo expression of murine p-			
CC	selectin.";			
CC	Blood 80:795-800(1992).			
CC	-I- FUNCTION: CA(2+)-DEPENDENT RECEPTOR FOR MYELOID CELLS THAT BINDS			
CC	TO CARBOHYDRATES ON NEUTROPHILS AND MONOCYTES. MEDIATES THE			
CC	INTERACTION OF ACTIVATED ENDOTHELIAL CELLS OR PLATELETS WITH			
CC	LEUCOCYTES. THE LIGAND RECOGNIZED IS SIALIN-LEWIS X.			
CC	-I- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.			
CC	-I- TISSUE SPECIFICITY: STORED IN THE ALPHA-GRANULES OF PLATELETS			
CC	AND WEIBEL-PALADE BODIES OF ENDOTHELIAL CELLS. UPON CELL			
CC	ACTIVATION BY AGONISTS, P-SELECTIN IS TRANSPORTED RAPIDLY TO			
CC	THE CELL SURFACE.			
CC	-I- INDUCTION: BY TUMOR NECROSIS FACTOR ALPHA.			

CC	-1- SIMILARITY: TO OTHER SELECTINS/LECAMs.	
CC	-1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.	
CC	-1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.	
CC	-1- SIMILARITY: CONTAINS 8 SUSHI (SCR) REPEATS; MOUSE P-LECTIN LACKS THE HUMAN SUSHI-2 EQUIVALENT.	
CC	-----	
CC	THIS SWISS-PROT ENTRY IS COPYRIGHTED. IT IS PRODUCED THROUGH A COLLABORATION	
CC	CC between the Swiss Institute of Bioinformatics and the EMBL Outstation	
CC	CC the European Bioinformatics Institute. There are no restrictions on its	
CC	CC use by non-profit institutions as long as its content is in no way	
CC	CC modified and this statement is not removed. Usage by and for commercial	
CC	CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/	
CC	CC or send an email to license@isb-sdb.ch).	
CC	-----	
DR	EMBL: M87861; AAA40008.1; -	
DR	EMBL: M72332; AAA37712.1; -	
DR	PIR: A42755; A42755.	
DR	HSSP: P16109; 1FSB.	
DR	MGI: 96280; SELP.	
DR	PRINTS: PR00343; SELECTIN.	
DR	PROSITE; PS00022; EGF_1; 1.	
DR	PROSITE; PS01186; EGF_2; 1.	
DR	PROSITE; PS00615; C-TYPE_LECTIN_1; 1.	
DR	PROSITE; PS00041; C-TYPE_LECTIN_2; 1.	
DR	PFAM; PF00008; EGF; 1.	
DR	PFAM; PF00059; lectin_c; 1.	
DR	Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;	
KW	selectin; signal; Sushi; Repeat.	
FT	SIGNAL	1
FT	CHAIN	42
FT	DOMAIN	42 768
FT	TRANSMEM	710 709
FT	DOMAIN	710 733
FT	DOMAIN	734 768
FT	DOMAIN	58 158
FT	DOMAIN	159 195
FT	DOMAIN	199 700
FT	REPEAT	199 258
FT	REPEAT	261 320
FT	REPEAT	323 382
FT	REPEAT	385 444
FT	REPEAT	447 506
FT	REPEAT	509 568
FT	REPEAT	579 638
FT	REPEAT	641 700
FT	DISULFID	60 158
FT	DISULFID	131 150
FT	DISULFID	163 174
FT	DISULFID	168 183
FT	DISULFID	185 194
FT	DISULFID	200 244
FT	DISULFID	230 257
FT	DISULFID	262 306
FT	DISULFID	292 319
FT	DISULFID	324 368
FT	DISULFID	354 381
FT	DISULFID	386 430
FT	DISULFID	416 443
FT	DISULFID	448 492
FT	DISULFID	478 505
FT	DISULFID	510 554
FT	DISULFID	540 567
FT	DISULFID	580 624
FT	DISULFID	610 637
FT	DISULFID	642 686
FT	DISULFID	672 699
FT	CARBOHYD	398 603
FT	CARBOHYD	603 693
FT	CARBOHYD	654 654
FT	CARBOHYD	661 661
FT	CARBOHYD	679 679
FT	SITE	756 759
FT	CONFLICT	724 724
FT	A -> E (IN REF. 2).	

```

SQ      SEQUENCE          768 AA;  83098 MW;  E5173074D266BE68 CRC64;
Query Match              13.5%, Score 343; DB 1; Length 768;
Best Local Similarity 26.9%; Pred. No. 9,54e-62;
Matches 77; Conservative 52; Mismatches 134; Indels 23; Gaps 23;

Db      FECOPGYRAGRSNTLHCTGSGOW-S-EPLPTTCEAIACEPP-EIPIHSMDCVPSGTFFCY 346
QY      :|::||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
32 YKCRPGRGSLGNVIIMCR-KGEWALNPLKKCQRCPGHGDYF-GTF-LTGDNVEY 88
Db      NSSCTFLAAGEFYLVKGN-DAIQACDSQWTAPAFCEALOCPEFPVPYSKAQ-VNCS-DPF 403
QY      :|::||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
89 GKAVAYTCNGEYOLLGEINRECDTDG-WINDIPICEIVACLPTATENGNIYSSANEPD 147
Db      GTLLY-QSVCSFCDBESLLVGASVIRCLATGHMNGAPPCQAVSCAPMLSPENSGMTCV 462
QY      :|::||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
148 REHFEGQAV-RFCNSGSYKLEGGDEMHCSDDGFWSKRKPCKVEILSCSP-DVINGSPIS- 204
Db      QPLGNSTYKSTCFPMCDGEHYLTSGPERLDSPSGHMTGPRTCEAIKCPIGAPEGNDL 522
QY      :|::||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
205 KLIITYKENERF-QYKNMGLEYSERDVACTBEG-WR-PLPSEBKSNDNPYLPN-GDIS 260
Db      CSVHAEFGVGSICHFSCNDEFELLSENVE-CYVSGRWSAAPPPTC 567
QY      :|::||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
261 PLIKHRTG-DEI-TYOCHRGFYPATRGTAKTCTSTG-W-IPAPRC 302

RESULT  14
ID      CCPH_HSVSA           STANDARD;             PRT;       360 AA.
AC      C001016;
DT      01-APR-1993 (rel. 25, Created)
DT      01-APR-1993 (rel. 25, Last sequence update)
DT      15-JUL-1998 (rel. 36, Last annotation update)
DE      COMPLEMENT CONTROL PROTEIN HOMOLOG PRECURSOR (CCPH).
GN      4 OR CCPH.
OS      Herpesvirus saimiri (strain 11).
OC      Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC      Gamaherpesvirinae.
RN      [1]
RX      ALBRECHT J.-C., NICHOLAS J., BILLER D., CAMERON K.R., BIESINGER B.,
RA      NEWMAN C., WITTMANN S., CRAXTON M.A., COLEMAN H., FLECKENSTEIN B.,
RA      HONESS R.W.;
RT      "Primary structure of the herpesvirus saimiri genome.";
RL      J. Virol. 66:5047-5058(1992).
RN      [2]
RP      SIMILARITY TO CCP.
RX      MEDLINE; 92260674.
RA      ALBRECHT J.-C., FLECKENSTEIN B.;
RT      "New member of the multigene family of complement control proteins in
        herpesvirus saimiri.";
RL      J. Virol. 66:3937-3940(1992).
CC      -1- ALTERNATIVE PRODUCTS: A MEMBRANE-BOUND FORM AND A SECRETED FORM
        ARE PRODUCED BY ALTERNATIVE SPLICING OF THE SAME GENE.
CC      -1- SIMILARITY: BELONGS TO THE SUPERFAMILY OF THE REGULATORS OF
        COMPLEMENT ACTIVATION (RCA).
CC      -1- SIMILARITY: CONTAINS 3 SUSHI (SCR) REPEATS.
CC      CC
CC      THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; X64346; CAA45626.1; -
DR      EMBL; X64346; CAA45627.1; -
DR      EMBL; X60283; CAA42823.1; -
DR      EMBL; X60283; CAA42822.1; -
DR      PIR; B42534; WWBEEZ.
DR      PIR; A42534; WWBEEI.

```


Query	Match	Similarity	Score	DB	Length	Best Local	Pred. No.	Indels	Gaps
Matches	74;	Conservative	45;	Mismatches	91;	Indels	22;	Gaps	17
Db	44	YPNGTTHVTVREGYAKRPVQVTVTCVNGM-TV-P-KKCKRKCKSTPDQLNGRVTYG-99	13.4%	Score 341;	DB 1;	Length 360;			
Qy	24	YPEGQATAYKCRPGYRSIGNVIMCRKGEWALNLPKCRKRGPHDPDTPFGFTILGG-83	31.9%	Pred. No. 3,31e-61;					
Db	100	NLY-IGSYITTTCSNGSYLIGST-SACLLKRGGVNDTPRPPICDIKKCP--PQIAN-155							
Qy	84	NVEFGVAKVAYTCNCGYLLGEINYNREC--DTDG--WTNDIPICEVVKCLPVAPENGK-138							
Db	156	GHTHTVWK-DF-VTYADTVTYSNDETKTLTLPGSSKLCSETGSWVNGETCEEFCKLP-213							
Qy	139	IYSSAMEPDRHHPQANRFVNCNSGYKE--GDEEMHCSDDGFW-SKSKPCVETSCSKSP-195							
Db	214	QVANAIVYEVKRSANISMOYLHINVKCYKGFMLYGETPNT-CNHGWSPAIPEC-264							
Qy	196	DVING-SPISQKIYKENRFPQYKCMNGYE-YSEKGDVATESGWRP-LPSC-244							
RESULT	15								
ID	F13B_HUMAN	STANDARD;	PRT;	661	AA.				
AC	P05160;								
DT	13-AUG-1987 (Rel. 05, Created)								
DT	01-AUG-1990 (Rel. 15, Last annotation update)								
DT	15-JUL-1999 (Rel. 38, Last annotation update)								
DE	COAGULATION FACTOR XII B CHAIN PRECURSOR (EC 2.3.2.13) (PROTEIN-GLUTAMINE GAMMA-GLUTAMYLTRANSFERASE B CHAIN) (TRANSGLUTAMINASE B CHAIN).								
DE	GLUTAMINE GAMMA-GLUTAMYLTRANSFERASE B CHAIN) (TRANSGLUTAMINASE B CHAIN).								
GN	F13B.								
OS	Homo sapiens (Human).								
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;								
OC	Euteria; Primates; Catarrhini; Homiidae; Homo.								
RN	[1]								
RP	SEQUENCE FROM N.A.								
RA	MEDLINE: 91105054.								
RT	Bottenus R.E., Ichinose A., Davie E.W.;								
RT	"Nucleotide sequence of the gene for the b subunit of human factor XIII."								
RT	Biochemistry 29:11195-11209(1990).								
RN	[2]								
RP	SEQUENCE OF 2-661 FROM N.A.								
RX	MEDLINE: 87026535.								

```

RA Ichinose A., McMullen B.A., Fujikawa K., Davie E.W.;  

RT "Amino acid sequence of the b subunit of human factor XIII, a protein  

RL Biochemistry 25:4633-4638(1986).  

RN [3]  

RP REVISIONS.  

RA Ichinose A.;  

RL Submitted (FEB-1987) to the EMBL/GenBank/DDBJ databases.  

RN [4]  

RC SEQUENCE OF 1-20 FROM N.A.  

RC TISSUE-LIVER;  

RX MEDLINE; 90251467.  

RA Grunmann U., Nerlich C., Rein T., Zettlmeissl G.;  

RT "Two genetic defects in a patient with complete deficiency of the b-  

RL subunit for coagulation factor XIII.";  

RL Blood 82:145-150(1993).  

CC -I- FUNCTION: THE B CHAIN OF FACTOR XIII IS NOT CATALYTICALLY ACTIVE,  

CC BUT IS THOUGHT TO STABILIZE THE A SUBUNITS AND REGULATE THE RATE  

CC OF TRANSGUTAMINASE FORMATION BY THROMBIN.  

CC -I- SUBUNIT: TETRAMER OF TWO A CHAINS AND TWO B CHAINS.  

CC -I- DISEASE: A DEFICIENCY IN FAL3 CAN RESULT IN A LIFE LONG BLEEDING  

CC TENDENCY, DECEPTIVE WOUND HEALING, AND HABITUAL ABORTION.  

CC -I- SIMILARITY: CONTAINS 10 SUSHI (SCR) REPEATS.  

-----  

CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  

CC the European Bioinformatics Institute. There are no restrictions on its  

CC use by non-profit institutions as long as its content is in no way  

CC modified and this statement is not removed. Usage by and for commercial  

CC entities requires a license agreement (See http://www.isb.ch/announce/  

CC or send an email to license@isb.ch).  

-----  

DR DR EMBL; M64554; AA51821.1; ALT_SEQ.  

DR EMBL; M14057; AA88042.1; -;  

DR EMBL; X51823; CA36123.1; -;  

DR PIR; A23830; A23830.  

DR PIR; A36397; A36397.  

DR PIR; S09980; S09980.  

DR HSSP; P08603; IHFI.  

DR MIN; J34580; -;  

DR PFAM; PF00084; sushi; 8.  

KW Transferrase; Plasma; Blood coagulation; Repeat; Glycoprotein; Signal;  

KW Sushi; Disease mutation.  

FT SIGNAL 1 20  

FT CHAIN 21 661  

FT DOMAIN 24 647  

FT REPEAT 24 88  

FT REPEAT 90 147  

FT REPEAT 152 209  

FT REPEAT 212 268  

FT REPEAT 273 328  

FT REPEAT 335 390  

FT REPEAT 395 451  

FT REPEAT 453 516  

FT REPEAT 523 579  

FT REPEAT 581 647  

FT DISULFID 25 76  

FT DISULFID 59 87  

FT DISULFID 91 135  

FT DISULFID 118 146  

FT DISULFID 153 197  

FT DISULFID 180 208  

FT DISULFID 213 255  

FT DISULFID 241 267  

FT DISULFID 274 316  

FT DISULFID 302 327  

FT DISULFID 336 378  


```

FT DISULFID 364 389 BY SIMILARITY.
 FT DISULFID 396 439 BY SIMILARITY.
 FT DISULFID 425 450 BY SIMILARITY.
 FT DISULFID 454 505 BY SIMILARITY.
 FT DISULFID 486 515 BY SIMILARITY.
 FT DISULFID 524 567 BY SIMILARITY.
 FT DISULFID 553 578 BY SIMILARITY.
 FT DISULFID 582 636 BY SIMILARITY.
 FT DISULFID 616 646 BY SIMILARITY.
 FT CARBOHYD 162 162 POTENTIAL.
 FT CARBOHYD 545 545 CELL ATTACHMENT SITE.
 FT SITE 617 619 C-> F (IN F13B DEFICIENCY).
 FT VARIANT 450 450 /Ftid=VAR_007475.
 SQ SEQUENCE 661 AA; 75491 MM; 57A2FB46560857F2 CRC64;

Query Match 13.4%; Score 340; DB 1; Length 661;
 Best Local Similarity 28.7%; Pred. No. 6.15e-61;
 Matches 89; Conservative 57; Mismatches 134; Indels 30; Gaps 24;

Db 178 YECATGYTTAGKRTKEVECLTYGW-SLTP--KCTKLKSSLRLENGYFH-PVKQYEE 233
 QY 32 YKCRPGY-RSLGN-VIMY-CRKGEWVALNPLKCGKRRPCGHGDTPTFTLTGQNVPEY 88
 Db 234 GDVVOFCHENYIYLSGS-DLIQCTNFGWPPSPVCEGRNRNCPPEPLPINSKIOTHT-T 291
 QY 89 GVKAVYTCNEGYOLLGEINRECDTGDWTNDIPICEV--VKCLPYTAPENGIIVSSAMEP 146
 Db 292 YR-H-GSIWIECELNFEIHGSAEIRC-EDGKWT-EPPKCIEGQEKVACEEPPFIENGA 346
 QY 147 DREYHFGQAVAFVYCNISGKIKIGDEMHCSDDGFMSKEKPKCEV---ISCKSPDYI-NGS 201
 Db 347 ANHSHKIYNGDKTYACKSGYLLHGSNEITCNRGKWTLPPECVENNENCKHPPVWNGA 406
 QY 202 PISQIITYKENERPOYCKNMGEYSEKDAVCTESGWRPLPSC-EEK-SCDNP-YIPNGD 258
 Db 407 VADGIILASVATGSSVEYRC-NEYI-LRGSKISRCEQGWSSPVC-LEPCT-VNVDYMN 462
 QY 259 YSP-LRIKHRTGDEITYQCRNGFYPATRGNT-AKCTSTGWLIPAPRCTLPCDYPDIKHGG 316
 Db 463 RNNIEMRWKY 472
 QY 317 LYHENMRRPY 326

Search completed: Thu Jun 8 21:46:19 2000
 Job time : 15 secs.

 MIPREH

 (TM)

Release 3.1a John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

MSPrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Thu Jun 8 21:46:38 2000; Maspar time 30.28 Seconds
 753.376 Million cell updates/sec

Tabular output not generated.

Title: >US-09-316-163-11
 Description: (1-329) from US09316163.pep
 Perfect Score: 2539
 Sequence: 1 EDGNELPPRRNTETILIGSMS.....PDIKHGGLYHNNRRPFV 329

Scoring table: PAM 150
 Gap 11

Searched: 228878 seqs; 80334122 residues

Post-processing: Minimum Match 0.8
 Listing first 45 summaries

Database:

sptrembl12
 1:sp.archaea 2:sp.bacteria 3:sp.fungi 4:sp.human
 5:sp.invertebrate 6:sp.mammal 7:sp.mmc 8:sp.organelle
 9:sp.phage 10:sp.plant 11:sp.rodent 12:sp.unclassified
 13:sp.vertebrate 14:sp.virus

Statistics: Mean 43.988; Variance 63.846; scale 0.689

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	2539	100.0	449	4	Q14570	COMPLEMENT FACTOR H PR	0.00e+00
2	1236	48.7	669	6	Q28085	CCP MODULES 3-12, WITH	0.00e+00
3	603	23.7	1053	13	Q91275	COMPLEMENT REGULATOR	1.99e-130
4	476	18.7	657	4	Q14006	COMPLEMENT H FACTOR (F	1.99e-95
5	438	17.3	452	11	Q1407	COMPLEMENT FACTOR H-RE	4.10e-85
6	436	17.2	645	14	Q9WRU2	COMPLEMENT BINDING PRO	9.20e-84
7	433	17.1	303	11	Q61405	COMPLEMENT FACTOR H-RE	1.10e-82
8	429	16.9	550	14	Q40912	ORF 04.	2.05e-82
9	428	16.9	550	14	P88903	ORF 04.	7.11e-82
10	426	16.8	522	6	Q28769	COMPLEMENT RECEPTOR (F	4.57e-81
11	423	16.7	2014	6	Q29530	COMPLEMENT RECEPTOR 1	2.33e-74
12	412	16.2	1911	6	Q29538	COMPLEMENT RECEPTOR 1	2.64e-77
13	409	16.1	2039	4	Q16745	COMPLEMENT RECEPTOR 1	2.64e-77
14	409	16.1	2489	4	Q16745	COMPLEMENT RECEPTOR 1	2.64e-77
15	403	15.9	661	6	Q29531	COMPLEMENT RECEPTOR 1	1.07e-75
16	398	15.7	259	14	P87616	41KBP FRAGMENT FROM LE	5.90e-72
17	389	15.3	522	6	Q28797	UNKNOWN PROTEIN (FRAGM	1.09e-71
18	388	15.3	560	5	Q22328	COSMID T07H6.	4.32e-70
19	382	15.0	360	14	Q8XT08	COMPLEMENT CONTROL PRO	5.00e-69
20	378	14.9	417	11	Q35520	512 ANTIGEN (FRAGMENT)	

21	372	14.7	315	6	Q28770	COMPLEMENT RECEPTOR (F	1.96e-67
22	373	14.7	679	11	Q99254	COMPLEMENT RECEPTOR TY	1.06e-67
23	373	14.7	740	4	Q95508	DJ780M13.1.2 (SELECTIN	1.06e-67
24	363	14.3	533	11	Q08569	ACROSOMAL MATRIX COMPO	4.72e-65
25	362	14.3	579	11	Q60736	ZONA PELLUCIDA 3 RECP	8.67e-65
26	360	14.2	808	11	Q61408	COMPLEMENT FACTOR H-RE	2.93e-64
27	357	14.1	263	14	Q07033	D12L PROTEIN.	1.81e-63
28	357	14.1	263	14	Q89859	HOMOLOG OF VACCINIA VI	1.81e-63
29	355	14.0	263	14	Q89076	CELL ADHESION MOLECULE	6.10e-63
30	354	13.9	754	6	Q28290	B19L.	1.12e-62
31	350	13.8	253	14	Q89061	D15L.	1.26e-61
32	343	13.5	974	5	P91658	FURROWED.	8.72e-60
33	341	13.4	559	11	Q63135	COMPLEMENT REGULATORY	2.92e-59
34	338	13.3	349	4	Q15429	CD46.	1.78e-58
35	336	13.2	497	11	Q63612	512 ANTIGEN PRECURSOR.	5.94e-58
36	336	13.2	740	4	Q95507	DJ780M13.1.1 (SELECTIN	5.94e-58
37	331	13.0	285	6	Q19126	MEMBRANE COPACITOR PROT	1.20e-56
38	331	13.0	285	6	Q19127	MEMBRANE COPACITOR PROT	1.20e-56
39	328	12.9	1045	6	P79138	MEMBRANE COPACITOR PROT	7.28e-56
40	328	12.9	1045	6	Q46545	MEMBRANE COPACITOR PROT	7.28e-56
41	323	12.7	285	6	Q19121	MEMBRANE COPACITOR PROT	1.46e-54
42	318	12.5	377	6	Q62838	MEMBRANE COPACITOR PROT	2.90e-53
43	317	12.5	378	6	Q62837	MEMBRANE COPACITOR PROT	5.27e-53
44	316	12.4	363	6	Q02839	PORCINE MEMBRANE COFAC	9.57e-53
45	316	12.4	1087	4	Q14212	EPSTEIN-BARR VIRUS COM	9.57e-53

ALIGNMENTS

RESULT ID	1	PRELIMINARY;	PRT;	449 AA.
AC	Q14570: P78435;			
DT	01-NOV-1996 (TREMURel. 01, Created)			
DT	01-NOV-1996 (TREMURel. 01, Last sequence update)			
DT	01-NOV-1999 (TREMURel. 12, Last annotation update)			
DE	COMPLEMENT FACTOR H PRECURSOR.			
GN	HF OR CFH.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;			
CC	Eukaryota; Primates; Catarrhini; Homnidae; Homo.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE; 88134059.			
RA	RIPACHE J., DAY A.J., HARRIS T.J.R., SIM R.B.;			
RT	"The complete amino acid sequence of human complement factor H.";			
RL	Biochem. J. 249:593-602(1988).			
RN	[2]			
RP	SEQUENCE OF 226-449 FROM N.A.			
RX	MEDLINE; 86169701.			
RA	KRISTENSEN T., WEISSEL R.A., TACK B.F.;			
RT	"Structural analysis of human complement protein H: homology with Cab			
RL	binding protein, beta 2-glycoprotein I, and the Ba fragment of B2.";			
RN	J. Immunol. 136:3407-3411(1986).			
RN	[3]			
RP	SEQUENCE OF 1-19 FROM N.A.			
RA	VIR D.P., WILLIAMS S.A.;			
RT	Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.			
RN	[4]			
RP	SEQUENCE OF 1-9 FROM N.A.			
RA	DOMINGUEZ O.;			
RT	Thesis (1993). Immunologia. Hospital Trias I Pujol, Spain.			
RN	EMBL; X07523; CA330403.1;			
DR	EMBL; M12383; AA52013.1;			
DR	EMBL; U56979; AB01987.1;			
DR	EMBL; Z29665; CA82763.1;			
DR	HSSP; P10998; IYVC.			
DR	PFAM; PF00084; sushi; 7.			
KW	Signal.			
FT	SIGNAL	1	18	POTENTIAL.
FT	CHAIN	19	449	POTENTIAL.
FT	SEQUENCE	449 AA;	51033 MM;	61231EIB CRC32;
SQ	Query Match	100.0%;	Score 2539;	DB 4; Length 449;

Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 329; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Db 19 EDCEMLPPRRMTTEITGSMQSYEGTQATYKCPGPRSGNVTMCRKGEWALNLR 78
QY 1 EDCEMLPPRRMTTEITGSMQSYEGTQATYKCPGPRSGNVTMCRKGEWALNLR 60
Db 79 KQKRRPGCHPDTPFGFTLLTGNVFEYGVAVYTCNGYOLLGEINRECDTGTMTDI 138
QY 61 KQKRRPGCHPDTPFGFTLLTGNVFEYGVAVYTCNGYOLLGEINRECDTGTMTDI 120
Db 139 PICVVCCLPTAPENGKIVSSAEPREYHFGQAVRVVCSGKIKESDEMHCSDGFW 198
QY 121 PICVVCCLPTAPENGKIVSSAEPREYHFGQAVRVVCSGKIKESDEMHCSDGFW 180
Db 199 SKERPKCVESICKSPDVINGSPISOKIITYKNERFQYKCNNGYSESGDAVCTESGWRP 258
QY 181 SKERPKCVESICKSPDVINGSPISOKIITYKNERFQYKCNNGYSESGDAVCTESGWRP 240
Db 259 LPSCSEKSCNPIYNDYSPILRIKHTGDEITYQCNGFYPATRGMTAKCTGTGWIAP 318
QY 241 LPSCSEKSCNPIYNDYSPILRIKHTGDEITYQCNGFYPATRGMTAKCTGTGWIAP 300
Db 319 RCTLKPCDYPDIKGGIYHEMRPFPV 347
QY 301 RCTLKPCDYPDIKGGIYHEMRPFPV 329

```

RESULT 2
ID Q28085; PRELIMINARY; PRT; 669 AA.
AC Q28085;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
DE CCP MODULES 3-12, WITH PARTS OF CCP 2 AND 13 (FRAGMENT).
OS Bos taurus (Bovine).
RT Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-LIVER;
RX MEDLINE: 96202005.
RA SOAMES C.J., DAY A.J., SIM R.B.;
RT "Prediction from sequence comparisons of residues of factor H involved
in the interaction with complement component C3b."
RL Biochem. J. 315:533-531(1996).
DR HMRP; X98697; CA67257.1; -.
DR HMRP; F10988; IYVC.
DR PFAM; PF00084; sushi; 11.
FT NON_TER 1
FT SEQUENCE 669 AA; 75683 MW; FAF0D174 CRC32;
SQ

Query Match 48.7%; Score 1236; DB 6; Length 669;
Best Local Similarity 61.4%; Pred. No. 0.00e+00;
Matches 153; Conservative 35; Mismatches 59; Indels 2; Gaps 1;

```

Db 7 LAEGQFEYGAIVYTCDEGYOMVGMNRECDTGTMTDIICEVKKLPTEPENGKI 66
QY 80 LTGVNFEYGVAVYTCNGYOLLGEINRECDTGTMTDIICEVKKLPTEPENGKI 139
Db 67 FSDALEPDEYTYGVGVFECSNGYMLDPPKQIHCSAGSVASAEPTKVEICKSPVIIN 126
QY 140 VSSAEPREYHFGQAVRVVCSGKIKESDEMHCSDGFWKPKVEICKSPDVIN 199
Db 127 GOAVLPKATYTCNENVOYRCAGFEYGGQDGTVCYTKSGMTPAFTCIETCDPPRIPI 186
QY 200 GSPISOKIITYKNERFQYKCNNGYSESGDAVCTESGWRPPLPSCSEKSCNPIY 259
Db 187 RPELSKYRGQATITECKGFEFPIRTGATRDGWPVPVPCAKPKSPYIKRGLY 246
QY 260 SPLRKHTGTGEITYQCNGFYPATRGMTAKCTGTGWIAPRCTLKPCDYPDIKGGIYH 319

```

Db 247 SY--RGYFP 253
QY 320 ENMRPFP 328

RESULT 3
ID Q91275; PRELIMINARY; PRT; 1053 AA.
AC Q91275;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
DE COMPLEMENT REGULATORY PLASMA PROTEIN.
OS Parabrax nebulifer (barred sand bass).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
OC Neopterygii; Teleostei; Euteleostei; Acanthopterygii; Percomorpha;
OC Perciformes; Percoidae; Serranidae; Parabrax.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-LIVER;
RX MEDLINE: 94318039.
RA DAHMEN A., KALDOH T., ZIPPEL P.F., GIGLI I.;
RT "Cloning and characterization of a cDNA representing a putative
complement-regulatory plasma protein from barred sand bass (Parabrax
nebulifer)."
RL Biochem. J. 301:391-397(1994).
DR EMBL; L21703; AAA92556.1; -.
DR HMRP; P08603; IHRP.
DR PFAM; PF00084; sushi; 16.
FT SEQUENCE 1053 AA; 117597 MW; 0DF68EDB CRC32;
SQ

Query Match 23.7%; Score 603; DB 13; Length 1053;
Best Local Similarity 34.1%; Pred. No. 1.99e-130;
Matches 101; Conservative 49; Mismatches 129; Indels 17; Gaps 13;

```

Db 47 EASYPCGROVRGCVNYG-S-GFFKLVCYEGKWTGGA--KCQPRSCGHPGDAOFADPHL 102
QY 21 DQTYPGTQATYKCRGPRGSLGNVTMCRKGEWALNLRKQKRRPGCHPDTPFGFTLL 80
Db 103 AEGNDFVGSKYVYTCQKGYOMVSRINRYCAEGMDGVVPCESQCC-PLIHDNNVOY 161
QY 81 TGVNFEYGVAVYTCNGYOLLGEINRECDTGTMTDIICEVKKLPTEPENGKI 140
Db 162 IGG-PE-EATFGVNVFESCKSRSEILDSPELYCDEMGAPKPKAKIAPPIEN 218
QY 141 SSAFEPRREYHFGQAVRVVCSGKIKESDEMHCSDGFWKPKVEICKSPDVIN 199
Db 219 GNVPAIRKYNVDVLAHECDRAFKHIDRPSTCIKQIKAEWSPPLCESTKRLTMDG 278
QY 200 GSPISOKIITYKNERFQYKCNNGYSESGDA-V--CTESGWRPPLPSCSEKSCNPIY 256
Db 279 TRYEPAYRNLFSPGELTLYVCARTSWISTPOETSVYTCQDNGESINPTQGVNC 334
QY 257 GDYSPL-RIKHTGTGEITYQC-RNGFYPATRGNT--AKCTSIG-WIPAPRCTLKPC 307

```

RESULT 4
ID Q14006; PRELIMINARY; PRT; 657 AA.
AC Q14006;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
DE COMPLEMENT H FACTOR (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominiidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 88025472.
RA DAY A.J., RIPOCHE J., LYONS A., MCINTOSH B., HARRIS T.J., SIM R.B.;
RT "Sequence analysis of a cDNA clone encoding the C-terminal end of
human complement factor H.";

RL Bioscl. Rep. 7:201-207(1987).
 DR EMBL; M17517; AAA52016.1; -.
 DR HSSP; P08603; 1HFH.
 DR PFM; PF00084; sushi; 11.
 FT NON_TER 1
 SQ SEQUENCE 657 AA; 74247 MW; F4AB5238 CRC32;

Query Match 18.7%; Score 476; DB 4; Length 657;
 Best Local Similarity 27.5%; Pred. No. 1,99e-95;
 Matches 86; Conservative 72; Mismatches 132; Indels 23; Gaps 22;

Db 296 CSQ-PROTEINSSSSQSSAHGKLTSTCEGFRISENETTCYMGKW-S-SP-PQ 351
 QY 3 CNEPLPRNTEILGWSWQ--TYPEGQAIKCRPRGHSLSGNVIMVCRKSGVNLNLRK 61
 Db 352 CEGLPCKRPEISHGVNAMSDS-YOYGEVYTKCFEGGIDGPAIAK-CLGKWSHP-P 408
 QY 62 COKRPGHGTPTGTFTLTGNGVFEYVKNVYTCNNEGTLGELINRECDGTNDIP 121
 Db 409 SCIKTDCLSLPSENA-IPGEEK-D-VYKAGEVYTYCAIYKMDASNTCINSR-WT 464
 QY 122 ICEVVKCLPVTAPENGKIVSSAMEPDRHFHGOAVRFVCSNGYKIEGDEHNCSDDGFW 181
 Db 465 -GRPCRTSCVNPPTVONATVSRQMSKYPSEGRVRYOCSPEMFGD-EVVKCLNGNW 522
 QY 182 KEKPCVETSC-KSPDVINGSPIS-QKIIYKENERFOYKCMNGE-YSERGDVCTESGW 238
 Db 523 TEPPOCKDSTGCGPPPIDNGDTFPLSYAVAPASSVEYOCQN-LYOLEGNKRTICRG 581
 QY 239 RLPLSCSEKS--CDNRY-IPNDYSLPAIK-HRTGDELTTCRNGFYATGNTAKCTST 294
 Db 582 QWSEPPKC-LHPC 593
 QY 295 GWIPARCTLKPC 307

RESULT 5
 ID 061407 PRELIMINARY; PRT; 452 AA.
 AC 061407;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
 DE COMPLEMENT FACTOR H-RELATED PROTEIN.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA VIK D.P., MUNOZ-CANOVES P., KOZONO H., MARTIN L.G., TACK B.F.,
 RA CHAPLIN D.D.;
 RL J. Biol. Chem. 0:0-0(0).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA VIK D.P., MUNOZ-CANOVES P., KOZONO H., MARTIN L.G., TACK B.F.,
 RA CHAPLIN D.D.;
 RL SUBMITTED (APR-1990) to the EMBL/GenBank/DBJ databases.
 DR EMBL; M29010; AAA37415.1; -.
 DR HSSP; P08603; 1HFI.
 DR PFM; PF00084; sushi; 7.
 SQ SEQUENCE 452 AA; 51602 MW; 45COB61 CRC32;

Query Match 17.3%; Score 438; DB 11; Length 452;
 Best Local Similarity 56.4%; Pred. No. 4,10e-85;
 Matches 53; Conservative 16; Mismatches 25; Indels 0; Gaps 0;

Db 15 TAWLSTAKEETGSPYILNGITPRRIIHSDEIRYECNFGFYPTGTSKCTPTG 74
 QY 236 SGWRPLSCSEKSCDNYPINSGDISPLRIKHRTGDELTTCRNGFYATGNTAKCTSTG 295
 Db 75 WIPVPRCTLKPCERPOKRYGLYEEESLRPNFV 108
 QY 296 WIPARCTLKPCDYPRDKHGLYHENMRPRFV 329

RESULT 6
 ID 09WRU2 PRELIMINARY; PRT; 645 AA.
 AC 09WRU2;
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
 DE COMPLEMENT BINDING PROTEIN.
 OS Macaca mulatta rhadinovirus 17577.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Gammaherpesvirinae; Rhadinovirus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 99174001.
 RA SEARLES R.P., BERGOUNM E.P., AXTHELM M.K., WONG S.W.;
 RT "Sequence and genomic analysis of a Rhesus macaque rhadinovirus with
 RT similarity to Kapoli's sarcoma associated herpesvirus/human
 RT herpesvirus 8.";
 RL J. Virol. 73:3040-3053(1999).
 DR EMBL; AF083501; AAD21332.1; -.
 SQ SEQUENCE 645 AA; 71526 MW; 60FB82D6 CRC32;

Query Match 17.2%; Score 436; DB 14; Length 645;
 Best Local Similarity 30.3%; Pred. No. 1,42e-84;
 Matches 72; Conservative 50; Mismatches 94; Indels 17; Gaps 14;

Db 337 EKYSGVAVELICRPQFTKMOSTVSECLSNGTWAPNA--KCHRRKCPDPELLNGEYI 394
 QY 22 QTYPEGQAIYKCRPRGHSLSGNVIM-C-RKGVNVLNLRKQKRGHGTPTGTFT 79
 Db 395 VTSGEDFAFKYGTINITYKCNQYOLGSMWRICMLKDDLTVMERKAPICDIEKCPQ 454
 QY 80 LTGG-VNFEYGVKAVYTCNNEGTLGELINRECDGTNDIPCEVVKCLPVTA 133
 Db 455 ITNGKY--HEVY-D-F-YQYIDTYFSCNRDFSIVGDMETTCISNT--MKNPFCSEITCS 509
 QY 134 PENKIVSAMPEPDRHFHGOAVRFVCSNGYKIEGDEHNCSDGFWSKKPCVETSC 193
 Db 510 APNIAHGKLTGSSSVYKYGASVYIGCEGFTLIGSEISTCKDSMDPLPTC 562
 QY 194 SPDVINGSPIS-QKIIYKENERFOYKCMNGEYSERGDVCTESGW-RPLPSC 244

RESULT 7
 ID 061405 PRELIMINARY; PRT; 303 AA.
 AC 061405;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
 DE COMPLEMENT FACTOR H-RELATED PROTEIN.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA VIK D.P., MUNOZ-CANOVES P., KOZONO H., MARTIN L.G., TACK B.F.,
 RA MEDLINE; 90153969.
 RA VIK D.P., MUNOZ-CANOVES P., KOZONO H., MARTIN L.G., TACK B.F.,
 RA CHAPLIN D.D.;
 RL "Identification and sequence analysis of four complement factor H-
 RL related transcripts in mouse liver.";
 RL J. Biol. Chem. 265:3193-3201(1990).
 DR EMBL; M29007; AAA37413.1; -.
 DR HSSP; P10998; 1VVC.
 DR PFM; PF00084; sushi; 4.
 SQ SEQUENCE 303 AA; 34498 MW; 39350FD1 CRC32;

Query Match 17.1%; Score 433; DB 11; Length 303;
 Best Local Similarity 55.3%; Pred. No. 9,20e-84;
 Matches 52; Conservative 17; Mismatches 25; Indels 0; Gaps 0;

Db 15 TAWLSTAKEETGSPYILNGITPRRIIHSDEIRYECNFGFYPTGTSKCTPTG 74
 QY 236 SGWRPLSCSEKSCDNYPINSGDISPLRIKHRTGDELTTCRNGFYATGNTAKCTSTG 295

Db 75 WIPVRLTKRCEPPEKRYGRLLYEESLRPNPV 108
 QY 296 WIPARCTLKPCDYPDIKHGLYHNMRRPFV 329

RESULT 8
 ID 040912 PRELIMINARY; PRT: 550 AA.
 AC 040912;
 DT 01-JAN-1998 (Tremblrel. 05, Created)
 DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)
 DT 01-NOV-1999 (Tremblrel. 12, Last annotation update)
 DE ORF 04.
 OS Kaposi's sarcoma-associated herpesvirus (KSHV) (Human herpesvirus 8).
 OC viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Gammaherpesvirinae; Rhadinovirus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 97296220.
 RA NEIPEL F., ALBRECHT J.C., FLECKENSTEIN B.;
 RT "Cell-homologous genes in the Kaposi's sarcoma-associated rhadinovirus
 RT human herpesvirus 8: determinants of its pathogenicity?";
 RL J. Virol. 71:4187-4192(1997).
 DR EMBL: U93872; AAB62602.1;
 DR HSSP: P10998; IVC.
 DR PFAM: PF00084; sushi; 4.
 DR SEQUENCE 550 AA; 60648 MW; 55808984 CRC32;
 SO

Query Match 16.9%; Score 429; DB 14; Length 550;
 Best Local Similarity 29.8%; Pred. No. 1.10e-82;
 Matches 82; Conservative 59; Mismatches 110; Indels 24; Gaps 20;

Db 52 RCRSGYTTARNITATCLOGGW-S-EPTATCKKSCPNPGEIONGKVIPIHGGDLATKG 109
 QY 33 KCRPGYRSLG-NVIMYC-RKGEVVALNPLRCKCRGCHGDPFGFTLTGG-NVFEYG 89

Db 110 ANISYCNBEGYFLVREYRYCMIGASGOMAMSSPPFEKERC-H-R-P--KINGDF 163
 QY 90 YKAVYTCNBYQLG-E-INRECDTDG---WTNDIPICEVVKCLPYTPENCKIYSSAM 144

Db 164 KPDKDYEEYNDVAFECNNGYTLVGPSTACAVNNWTSMPTCELAGCKFSTYHGYPI 223
 QY 145 EPDREYH-FGQAVRFVCSNGYKIEGDEMHCSDDGWSKEKPCVEISCKSPDIVNGSPI 203

Db 224 QGFSLYTKKQSVTFACNDGFEVLKRGSPITTCNVTEMDPLPCVLEDIDDPNNSNGRLH 283
 QY 204 SQ-KIITYENEFQYCKMNGYFYSERGDVCTESGW-RPLPSCIEKSCDNPYIPN-GDYS 260

Db 284 PTP-NEKPNGNV-FQ-RSNYTEPTKPEDTHAAT 315
 QY 261 PLRIKRTGDELTTCORNGFY-PATRGNTAKCTST 294

RESULT 9
 ID P88903 PRELIMINARY; PRT: 550 AA.
 AC P88903;
 DT 01-MAY-1997 (Tremblrel. 03, Created)
 DT 01-MAY-1997 (Tremblrel. 03, Last sequence update)
 DT 01-NOV-1999 (Tremblrel. 12, Last annotation update)
 DE ORF 4.
 OS Kaposi's sarcoma-associated herpesvirus (KSHV) (Human herpesvirus 8).
 OC viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Gammaherpesvirinae; Rhadinovirus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 97121480.
 RA RUSSO J.J., BOHENZKY R.A., CHEN M.C., CHEN J., YAN M., MADDALENA D.,
 RA PARRY J.P., PERUZZI D., EDELMAN I.S., CHANG Y., MOORE P.S.;
 RT "Nucleotide sequence of the Kaposi sarcoma-associated herpesvirus
 RT (HHV8).";
 RL Proc. Natl. Acad. Sci. U.S.A. 93:14862-14867(1996).
 DR SEQUENCE FROM N.A.

RX MEDLINE: 97094384.
 RA MOORE P.S., BASHOFF C., WEISS R.A., CHANG Y.;
 RT "Molecular mimicry of human cytokine and cytokine response pathway
 RT genes by KSHV";
 RL Science 274:1739-1744(1996).
 DR EMBL: U75698; AAC57082.1;
 DR HSSP: P10998; IVC.
 DR PFAM: PF00084; sushi; 4.
 DR SEQUENCE 550 AA; 60687 MW; CEAZCABC CRC32;
 SO

Query Match 16.9%; Score 428; DB 14; Length 550;
 Best Local Similarity 29.8%; Pred. No. 2.05e-82;
 Matches 82; Conservative 59; Mismatches 110; Indels 24; Gaps 20;

Db 52 RCRSGYTTARNITATCLOGGW-S-EPTATCKKSCPNPGEIONGKVIPIHGGDLATKG 109
 QY 33 KCRPGYRSLG-NVIMYC-RKGEVVALNPLRCKCRGCHGDPFGFTLTGG-NVFEYG 89

Db 110 ANISYCNBEGYFLVREYRYCMIGASGOMAMSSPPFEKERC-H-R-P--KINGDF 163
 QY 90 YKAVYTCNBYQLG-E-INRECDTDG---WTNDIPICEVVKCLPYTPENCKIYSSAM 144

Db 164 KPDKDYEEYNDVAFECNNGYTLVGPSTACAVNNWTSMPTCELAGCKFSTYHGYPI 223
 QY 145 EPDREYH-FGQAVRFVCSNGYKIEGDEMHCSDDGWSKEKPCVEISCKSPDIVNGSPI 203

Db 224 QGFSLYTKKQSVTFACNDGFEVLKRGSPITTCNVTEMDPLPCVLEDIDDPNNSNGRLH 283
 QY 204 SQ-KIITYENEFQYCKMNGYFYSERGDVCTESGW-RPLPSCIEKSCDNPYIPN-GDYS 260

Db 284 PTP-NEKPNGNV-FQ-RSNYTEPTKPEDTHAAT 315
 QY 261 PLRIKRTGDELTTCORNGFY-PATRGNTAKCTST 294

RESULT 10
 ID Q28769 PRELIMINARY; PRT: 522 AA.
 AC Q28769;
 DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
 DT 01-NOV-1999 (Tremblrel. 12, Last annotation update)
 DE COMPLEMENT RECEPTOR (FRAGMENT).
 OS Papio hamadryas (Hamadryas baboon).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Cercopithecoidea; Cercopithecoinae;
 OC Papio.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX TISSUE-BONE MARROW;
 RA BIRMINGHAM D.J., LOGAR C.M., SHEN X.P., CHEN W.;
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL: L77977; AAA99004.1;
 DR HSSP: P08603; IHFI.
 DR PFAM: PF00084; sushi; 7.
 DR NON_TER 1
 DR SEQUENCE 522 AA; 56626 MW; 7862072C CRC32;
 SO

Query Match 16.8%; Score 426; DB 6; Length 522;
 Best Local Similarity 32.1%; Pred. No. 7.11e-82;
 Matches 11; Conservative 74; Mismatches 119; Indels 42; Gaps 34;

Db 52 EQLPAPPTN-LIDA-SE--FVPGTYLYKCELPYHGKRPST-ICUKNSVWTSKAD--KC 104
 QY 4 NELPPRRNTIELTGSWSDQTYPEGTQAIYCRGYSGLGVIMWC-RKGEVVALNPLRKC 62

Db 105 TRKSCRNPKC-PVNGAVHVKDI-OFQSQIYNSCNGYRILGSSATCTIGNTVWIND 162
 QY 63 QKRPCGPGTPTGTTLGGVNFETGVKAVYTCNBYQLG-E-INRECDTDGNTND 119

Db 163 TPICEIIPGCLPTIA-NGDFISTRE--YFPYGSVYVYRCLGSGRKLFLPEPST 218
 QY 120 IPICEVVK-LPVTADENGKIYSSAMEPDBREYHFGQAVRFVGN--SGYK-I-E--GDEEM 172

Query	Subject	Score	DB	Length	Year
1391	FPVGTSLATNTECRPG--FEKMTSISGLEUWVSSVED--NCRRSKSGPPE-PFGMWI	16.7%	DB 6	1445	2014
24	YPGSTAIKCRGPGKSLGNVIMV-CRNG-EWVALNPLKCKRCPGHPGTPTFLT	Best local similarity 29.1%;	Pred. No. 4,578-81;		
	Matches 95; Conserved 74; Mismatches 119; Indels 39; Gaps 28				
1391	FPVGTSLATNTECRPG--FEKMTSISGLEUWVSSVED--NCRRSKSGPPE-PFGMWI	16.7%	DB 6	1445	2014
24	YPGSTAIKCRGPGKSLGNVIMV-CRNG-EWVALNPLKCKRCPGHPGTPTFLT	Best local similarity 29.1%;	Pred. No. 4,578-81;		
	Matches 95; Conserved 74; Mismatches 119; Indels 39; Gaps 28				
1446	NTPT-OFGSTVANSNGEGRLLGSPSTTCLVSGNNVYTDKKAIPICEIISCEPPTISNGD	16.7%	DB 6	1504	2014
82	GNVFEYGVKAVYTCNEGVLGGE--INVRCDTD-GWINDIPICEVAKCLPVAPENGK	Best local similarity 29.1%;	Pred. No. 4,578-81;		
	Matches 95; Conserved 74; Mismatches 119; Indels 39; Gaps 28				
1505	FYSNNRAS---FNGCVIYVYQCHTGTDGSQLFELVGERIYCTISKDDQGVWSSPPRCI	16.7%	DB 6	1561	2014
139	IYSSAMEPDREYFEFGAVRVCNSG-----YKIBDEEMHC-S-DD--GFWSKKPKCV	Best local similarity 29.1%;	Pred. No. 4,578-81;		
	Matches 95; Conserved 74; Mismatches 119; Indels 39; Gaps 28				
1562	STNKCAPAEYENAIRPGNRSFSLLEIRFCQDPGFVWVGSHTVQCOTNGRWGKLPHC	16.7%	DB 6	1621	2014
189	EIS-CNSPDVINSPTS-OXIIYKEMERQYKCMNGEYSESGDAVCTBSG-WRP-LPSC	Best local similarity 29.1%;	Pred. No. 4,578-81;		
	Matches 95; Conserved 74; Mismatches 119; Indels 39; Gaps 28				
1622	S-RVCCDPPEILGHTPSHQDNFSQGVYFVSCPG-Y-DLRGASLHCTQOGWMSPEA	16.7%	DB 6	1678	2014
245	EESKCDNP-YIPNGDYSPLRIKH-RTGDELTYOCCRNGFYPATRGNTA-KCTSTG-WIP-A	Best local similarity 29.1%;	Pred. No. 4,578-81;		
	Matches 95; Conserved 74; Mismatches 119; Indels 39; Gaps 28				
1679	PRCTVAKCDDELGLPHGRVLPPLNQ 1705	Best local similarity 29.1%;	Pred. No. 4,578-81;		
	Matches 95; Conserved 74; Mismatches 119; Indels 39; Gaps 28				
300	PRCTLAKPCD-Y-PDIKHG-GLYHENMR 323	Best local similarity 29.1%;	Pred. No. 4,578-81;		
	Matches 95; Conserved 74; Mismatches 119; Indels 39; Gaps 28				

[illegible]

DR EMBL: L17394; AAB60694.1; JOINED.
 DR EMBL: L17395; AAB60694.1; JOINED.
 DR EMBL: L17396; AAB60694.1; JOINED.
 DR EMBL: L17397; AAB60694.1; JOINED.
 DR EMBL: L17398; AAB60694.1; JOINED.
 DR EMBL: L17399; AAB60694.1; JOINED.
 DR EMBL: L17400; AAB60694.1; JOINED.
 DR EMBL: L17401; AAB60694.1; JOINED.
 DR EMBL: L17402; AAB60694.1; JOINED.
 DR EMBL: L17403; AAB60694.1; JOINED.
 DR EMBL: L17404; AAB60694.1; JOINED.
 DR EMBL: L17405; AAB60694.1; JOINED.
 DR EMBL: L17406; AAB60694.1; JOINED.
 DR EMBL: L17407; AAB60694.1; JOINED.
 DR EMBL: L17408; AAB60694.1; JOINED.
 DR EMBL: L17409; AAB60694.1; JOINED.
 DR EMBL: L17410; AAB60694.1; JOINED.
 DR EMBL: L17411; AAB60694.1; JOINED.
 DR EMBL: L17412; AAB60694.1; JOINED.
 DR EMBL: L17413; AAB60694.1; JOINED.
 DR EMBL: L17414; AAB60694.1; JOINED.
 DR EMBL: L17415; AAB60694.1; JOINED.
 DR EMBL: L17416; AAB60694.1; JOINED.
 DR EMBL: L17417; AAB60694.1; JOINED.
 DR EMBL: L17418; AAB60694.1; JOINED.
 DR EMBL: L17419; AAB60694.1; JOINED.
 DR EMBL: L17420; AAB60694.1; JOINED.
 DR EMBL: L17421; AAB60694.1; JOINED.
 DR EMBL: L17422; AAB60694.1; JOINED.
 DR EMBL: L17423; AAB60694.1; JOINED.
 DR HSSP: P08603; IHFI.
 DR PFAM: PF00084; sush1, 30.
 DR SEQUENCE 2039 AA; 223603 MW; A4972215 CRC32;

Query Match 16.18; Score 409; DB 4; Length 2039;
 Best Local Similarity 31.48; Pred. No. 2,64e-77;
 Matches 97; Conservative 66; Mismatches 110; Indels 36; Gaps 30;

DB 63 FPIGTLYNECRPGYSGRPSI-ICLNKSVWTKAKD-R-CRRKSCRRPPDPVNGMAYVIR 119
 QY 24 YPEGTQAIYKCRPGYSGRPSI-ICLNKSVWTKAKD-R-CRRKSCRRPPDPVNGMAYVIR 82
 DB 120 G-I-OFGSQIKYCTKGRYRLIGSSSATCIISGDTVIMNDETPICDRIPCGLPPTIT-NGD 176
 QY 83 GNVFEYGVKAVYTCNEGYYOLGE--INPREC-DTIDGTNDIPICEVYKC-LPTVAPENK 138
 DB 177 FISTNRE-N--FHYGSVYTYRCNPGSGKRVFELVEPSIYCTSDNDQVIGWGPAPQCI 233
 QY 139 IVSSAMEPDRHFGQAVRFVNCNGY--KI-E--GDEEMHC-S-DD--GFWSKSKPKCV 188
 DB 234 IPNCTPPNVEGILVSDNRLFSLEVEVEFRCPGPFVYKGRPRVYKQALNKEPELPS 293
 QY 189 EIS-CKSPDIVNGSPISQ-KIITYKENRFOYKCMNGEYSESGDAVCTE-SGWRP-LPSC 244
 DB 294 S-RVCOOPRDLVHARTORDKDNFSPGOEVRYSCEPG-Y-DIRGASAKRCPPOGDWSPAA 350
 QY 245 EEKSCDND-YIPNGYSPL-RIKHRTGDEITYQCRNGFYPAIRGNTA-KCISTG-WIDA- 299
 DB 351 PTCVKSQCD 359
 QY 300 PCTLKPCD 308

RESULT 14
 ID Q16744; PRELIMINARY; PRT: 2489 AA.
 AC Q16744;
 DT 01-NOV-1996 (Tremblrel 01, Created)
 DT 01-NOV-1996 (Tremblrel 01, Last sequence update)
 DT 01-NOV-1999 (Tremblrel 12, Last annotation update)
 DE COMPLEMENT RECEPTOR 1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 RC [1]

RP SEQUENCE FROM N.A.
 RX MEDLINE; 94065173.
 RA VIK D.P., WONG W.W.;
 RT Structure of the gene for the F allele of complement receptor type 1
 and sequence of the coding region unique to the S allele.*;
 RL J. Immunol. 151:6214-6224(1993).
 RN [2]

RP SEQUENCE FROM N.A.

RA VIK D.P., WONG W.W.;
 RL Submitted (JUN-1993) to the EMBL/GenBank/DBJ databases.

DR EMBL: L17418; AAB60695.1; JOINED.
 DR EMBL: L17390; AAB60695.1; JOINED.
 DR EMBL: L17391; AAB60695.1; JOINED.
 DR EMBL: L17392; AAB60695.1; JOINED.
 DR EMBL: L17393; AAB60695.1; JOINED.
 DR EMBL: L17394; AAB60695.1; JOINED.
 DR EMBL: L17395; AAB60695.1; JOINED.
 DR EMBL: L17396; AAB60695.1; JOINED.
 DR EMBL: L17397; AAB60695.1; JOINED.
 DR EMBL: L17398; AAB60695.1; JOINED.
 DR EMBL: L17399; AAB60695.1; JOINED.
 DR EMBL: L17400; AAB60695.1; JOINED.
 DR EMBL: L17401; AAB60695.1; JOINED.
 DR EMBL: L17402; AAB60695.1; JOINED.
 DR EMBL: L17403; AAB60695.1; JOINED.
 DR EMBL: L17404; AAB60695.1; JOINED.
 DR EMBL: L17405; AAB60695.1; JOINED.
 DR EMBL: L17406; AAB60695.1; JOINED.
 DR EMBL: L17407; AAB60695.1; JOINED.
 DR EMBL: L17408; AAB60695.1; JOINED.
 DR EMBL: L17409; AAB60695.1; JOINED.
 DR EMBL: L17410; AAB60695.1; JOINED.
 DR EMBL: L17411; AAB60695.1; JOINED.
 DR EMBL: L17412; AAB60695.1; JOINED.
 DR EMBL: L17413; AAB60695.1; JOINED.
 DR EMBL: L17414; AAB60695.1; JOINED.
 DR EMBL: L17415; AAB60695.1; JOINED.
 DR EMBL: L17416; AAB60695.1; JOINED.
 DR EMBL: L17417; AAB60695.1; JOINED.
 DR EMBL: L17418; AAB60695.1; JOINED.
 DR EMBL: L17419; AAB60695.1; JOINED.
 DR EMBL: L17420; AAB60695.1; JOINED.
 DR EMBL: L17421; AAB60695.1; JOINED.
 DR EMBL: L17422; AAB60695.1; JOINED.
 DR EMBL: L17423; AAB60695.1; JOINED.
 DR EMBL: L17424; AAB60695.1; JOINED.
 DR EMBL: L17425; AAB60695.1; JOINED.
 DR EMBL: L17426; AAB60695.1; JOINED.
 DR EMBL: L17427; AAB60695.1; JOINED.
 DR EMBL: L17428; AAB60695.1; JOINED.
 DR EMBL: L17429; AAB60695.1; JOINED.
 DR EMBL: L17430; AAB60695.1; JOINED.
 DR HSSP: P08603; IHFI.
 DR PFAM: PF00084; sush1, 37.
 DR SEQUENCE 2489 AA; 272846 MW; 5869B6F9 CRC32;

Query Match 16.18; Score 409; DB 4; Length 2489;
 Best Local Similarity 31.48; Pred. No. 2,64e-77;
 Matches 97; Conservative 66; Mismatches 110; Indels 36; Gaps 30;

DB 63 FPIGTLYNECRPGYSGRPSI-ICLNKSVWTKAKD-R-CRRKSCRRPPDPVNGMAYVIR 119
 QY 24 YPEGTQAIYKCRPGYSGRPSI-ICLNKSVWTKAKD-R-CRRKSCRRPPDPVNGMAYVIR 82
 DB 120 G-I-OFGSQIKYCTKGRYRLIGSSSATCIISGDTVIMNDETPICDRIPCGLPPTIT-NGD 176
 QY 83 GNVFEYGVKAVYTCNEGYYOLGE--INPREC-DTIDGTNDIPICEVYKC-LPTVAPENK 138
 DB 177 FISTNRE-N--FHYGSVYTYRCNPGSGKRVFELVEPSIYCTSDNDQVIGWGPAPQCI 233
 QY 139 IVSSAMEPDRHFGQAVRFVNCNGY--KI-E--GDEEMHC-S-DD--GFWSKSKPKCV 188
 DB 234 IPNCTPPNVEGILVSDNRLFSLEVEVEFRCPGPFVYKGRPRVYKQALNKEPELPS 293
 QY 189 EIS-CKSPDIVNGSPISQ-KIITYKENRFOYKCMNGEYSESGDAVCTE-SGWRP-LPSC 244

Db 294 S-RVCGPPDYVLAERTORDNFSPOGEVYSCPEG-Y-DLRGAASMRCTPOGWSPPA 350
QY 245 EKSQCDNP-YIPNGDYSPL-RIKHRTGDEITYOCRCNGFPATRGNTA-KCTSTG-WIPA- 299
Db 351 PTCEVKSQD 359
QY 300 PRCITLKPCD 308

RESULT 15
ID 029531 PRELIMINARY; PRT; 661 AA.
AC 029531.
DT 01-NOV-1996 (TIREMBLrel. 01, Created)
DT 01-NOV-1996 (TIREMBLrel. 01, Last sequence update)
DT 01-NOV-1999 (TIREMBLrel. 12, Last annotation update)
DE COMPLEMENT RECEPTOR 1 PRECURSOR (FRAGMENT).
GN CRI.
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Homiidae; Pan.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 94292799.
RA BIRMINGHAM D.J., SHEN X.P., HOURCADE D., NICKELLS M.W., ATKINSON J.P.;
RT "Primary sequence of an alternatively spliced form of CRI. Candidate
for the 75,000 Mr complement receptor expressed on chimpanzee
erythrocytes."
RT J. Immunol. 153:691-700(1994).
DR EMBL; L24921; AAA51439.1; -.
DR HSSP; P10998; IYVC.
DR PIRAM; PR00084; susht; 9.
KW Signal; Alternative splicing.
FT NON_TER 1
FT SIGNAL <1 16 POTENTIAL.
FT CHAIN 17 >661 COMPLEMENT RECEPTOR 1.
FT NON_TER 661
SQ SEQUENCE 661 AA; 72966 MW; 9D78E262 CRC32;

Query Match 15.98; Score 403; DB 6; Length 661;
Best Local Similarity 31.18; Pred. No. 1.07e-75;
Matches 96; Conservative 66; Mismatches 111; Indels 36; Gaps 30;

Db 38 FPIGTIYNTCRGGYGRPSI-ICLNKSVWTKAD-R-CRRKSCRNPPD-PVNGWVHYI 93
QY 24 YPEGTQAIYICRPGYRSLGNVIMVC-RKGEVALNPLKRCQKRCGHPGDTPTGFTTLTG 82
Db 94 KDI-QFGSQIKYSCITKGYRLIGSSSATCIIISDPTVIWNETPICDRIPCGLPPTIT-NGD 151
QY 83 GNVEFYGVKAVTCNDEYQLLGE-INYREC-DTDGWTNDIPICEVYKC-LPYTAPENCK 138
Db 152 FISTNRE-N-FHIGSVVYTRCNPBGSGKRVFELVGEPSIYCTSNDDQVGIWSPAPOCI 208
QY 139 IVSSAMEPDREHYFGQAVRFVCSNGY--KI-E-GDEEMHC-S-DD-GFWSKERPKCV 188
Db 209 IPNKCPTPNENGLVSDNLSFLSNVEVEFCOPGFVMKGPVRYCOALNKNEPELPSC 268
QY 189 EIS-CKSPDIVNSPISQ-KIYKENERFOYKCMGIEYERGDVCTE-SGWRP-LPSC 244
Db 269 S-RVCGPPDYVLAERTORDNFSPOGEVYSCPEG-Y-DLRGAASLRCTPOGWSPPAT 325
QY 245 EKSQCDNP-YIPNGDYSPL-RIKHRTGDEITYOCRCNGFPATRGNTA-KCTSTG-WIPA- 299
Db 326 PTCEVKSQD 334
QY 300 PRCITLKPCD 308

Search completed: Thu Jun 8 21:47:12 2000
Job time : 34 secs.

 MIPRELI (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

Mpsrch_p protein - protein database search, using Smith-Waterman algorithm
 Run on: Thu Jun 8 21:49:10 2000: Maspar time 13.73 Seconds
 Tabular output not generated. 738.315 Million cell updates/sec

Title: >US-09-316-163-14
 Description: (1-428) from US09316163.pep
 Perfect Score: 3371
 Sequence: 1 EDCKGPPRENSILSGMS.....DYTCENGSPPPKVRIRK 428

Scoring table: PAM 150
 Gap 11

Searched: 188963 seqs/23686106 residues

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database: a-gene36
 1:gene36p

Statistics: Mean 34.511; Variance 134.258; scale 0.257

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description	Pred. No.
1	1385	41.1	240	1 W39154	Human partial Complement	1.86e-134
2	1276	37.9	216	1 W39155	Clone PRB89FH410 CFH r	1.44e-122
3	712	21.1	578	1 Y09065	Human complement factor	1.36e-61
4	537	15.9	561	1 R13490	Human C4 binding prote	5.15e-43
5	445	13.2	543	1 R28557	CR1-4 (99H, 103E) anal	2.21e-33
6	445	13.2	543	1 R28552	CR1-4 (78T, 79D) analo	2.21e-33
7	446	13.2	720	1 W73147	Amino acid sequence of	1.74e-33
8	446	13.2	1930	1 W45899	Human complement recep	1.74e-33
9	446	13.2	2317	1 P92219	CR1 protein	1.74e-33
10	441	13.1	543	1 R28553	CR1-4 (85R, 87N) analo	5.77e-33
11	439	13.0	543	1 R28550	CR1-4 (114S) analogue.	9.31e-33
12	439	13.0	543	1 R28558	CR1-4 (109N, 110A, 111	9.31e-33
13	438	13.0	543	1 R28571	CR1-4 (364-367 NAAH) a	1.18e-32
14	437	13.0	543	1 R28550	CR1-4 (64K) analogue.	1.50e-32
15	439	13.0	1537	1 R11982	Partial human compleme	9.31e-33
16	439	13.0	2039	1 R36743	CR1.	9.31e-33
17	437	13.0	2039	1 R11810	Human complement type	1.50e-32
18	436	12.9	543	1 R28547	CR1-4 (52S, 53S, 54P)	1.91e-32
19	436	12.9	543	1 R28569	CR1-4 (36S-37E STRPITC	1.91e-32
20	436	12.9	543	1 R28558	CR1-4 (347T, 349Y) ana	1.91e-32
21	435	12.9	543	1 R28555	CR1-4 (92T) analogue.	2.42e-32
22	434	12.9	543	1 R28565	CR1-4 (121Q) analogue.	3.08e-32
23	433	12.8	543	1 R28566	CR1-4 (318R, 319N) ana	3.91e-32

24	433	12.8	543	1 R28570	CR1-4 (266-274 KUKTQTN	3.91e-32
25	432	12.8	543	1 R28549	CR1-4 (64K, 65T) analo	4.96e-32
26	432	12.8	543	1 R28563	CR1-4 (117T) analogue.	4.96e-32
27	431	12.8	543	1 R28561	CR1-4 (115T) analogue.	6.30e-32
28	431	12.8	543	1 R28551	CR1-4 (65T) analogue.	6.30e-32
29	430	12.8	543	1 R28545	CR1-4 (37Y) analogue.	8.01e-32
30	430	12.8	543	1 R28562	CR1-4 (116K) analogue.	8.01e-32
31	428	12.7	543	1 R28544	CR1-4 (35E) analogue.	1.29e-31
32	427	12.7	543	1 R28548	CR1-4 (57V, 59K) analo	1.64e-31
33	426	12.6	543	1 R28564	CR1-4 (116K, 117P) ana	2.08e-31
34	425	12.6	543	1 R28567	CR1-4 (318-321 RNP) a	2.64e-31
35	424	12.6	543	1 R28556	CR1-4 (94H) analogue.	3.35e-31
36	423	12.5	543	1 R28554	CR1-4 (92T, 94H) analo	4.26e-31
37	422	12.5	543	1 R28559	CR1-4 (114-117STRP, 12	5.41e-31
38	422	12.5	543	1 R28543	CR1-4 (35E, 37Y) analo	5.41e-31
39	416	12.3	543	1 R28546	CR1-4 (44T, 47D, 49L)	2.26e-30
40	411	12.2	579	1 W39924	Amino acid sequence of	7.44e-30
41	390	11.6	830	1 R65216	p-selectin.	1.09e-27
42	389	11.5	577	1 W06882	Membrane co-factor pro	1.38e-27
43	379	11.2	263	1 P92003	Deduced sequence of co	1.48e-26
44	363	10.8	254	1 R47154	Sequence of soluble co	6.46e-25
45	363	10.8	254	1 R47155	Sequence of soluble co	6.46e-25

ALIGNMENTS

RESULT 1
 ID W39154 standard; Protein: 240 AA.

AC W39154;
 DE 27-APR-1998 (first entry)

DE Human partial Complement factor H protein fragment 1.

KW Complement factor H; tumour associated antigen; renal cancer;

KM urogenital cancer; medicament; modulator.

OS Homo sapiens.

PN MO9738136-A1.

PD 16-OCT-1997.

PE 09-APR-1997; US-05710.

PF 06-MAR-1997; US-812481.

PR 09-APR-1996; US-015083.

PR 09-APR-1996; US-630048.

PR 06-MAR-1997; US-038614.

PA (BARD-) BARD DIAGNOSTIC SCI INC.

PI Enfield DL, Hase GM, Kinders RJ;

DR WPI: 97-512742/47.

DR N-PSDB: V02790.

PT Treating or screening for cancer, e.g. renal or urogenital cancer -

PT by modulating or detecting tumour associated human complement factor

PS H related antigen, or nucleic acid encoding it

PS Example 6B: Fig 6B: 104pp: English.

CC This partial protein sequence represents a region of the human

CC tumour associated complement factor H (CFH). This sequence is used

CC in the identification of complement factor H related proteins and

CC antigens isolated from clone PRB89FH410 (see W39155). The detection of

CC such proteins and a CFH antigens can be used in screening or for the

CC treatment of renal or urogenital cancer, e.g. bladder, cervical or

CC prostate cancer. Agents that may modulate this antigen could be used in

CC the manufacture of a medicament for the treatment of a tumour cell.

CC Sequence 240 AA;

CC

CC

CC

CC

CC

CC

CC

CC

```

QY 198 ENGDDGLKPKYKENERFOYKCKGQFVYKERGDVATCGSGWNPQSCHEMTCLTFPIPNQ 257
Db 181 DYSPLRIKHRTDEITTYOCNRNGFYPATRGNAKCTSTGWIAPRCLTKPCDPYDITKHGGL 240
QY 258 IYTHRIKHRIIDDEIRYECCKNGFPATRSVSKCTITGWIAPRCLTKPCDPFOPKHGRL 317

RESULT 2
ID W39155 standard; Protein; 216 AA.
AC W39155;
DT 27-APR-1998 (first entry)
DE Clone pRRB9FH410 CFH related protein fragment.
KM Complement factor H; tumour associated antigen; renal cancer;
OS urogenital cancer; medicament; modulator.
PN W09738136-A1.
PD 16-OCT-1997.
PF 09-APR-1997; U05710.
PR 06-MAR-1997; US-812481.
PR 09-APR-1996; US-015083.
PR 09-APR-1996; US-630048.
PR 06-MAR-1997; US-038614.
PA (BARD-) BARD DIAGNOSTIC SCI INC.
PI Enfield DL, Hase GM, Kinders RJ;
DR WPI: 97-512742/47.
DR N-PSDB: V02791.

PT Treating or screening for cancer, e.g. renal or urogenital cancer -
PT by modulating or detecting tumour associated human complement factor
PT H related antigen, or nucleic acid encoding it
PS Example 6B; Fig 6B; 104pp; English.
CC This partial protein is found in clone pRRB9FH410 and represents a
CC human tumour associated complement protein which homologs to a region of the
CC protein and a CFH antigen can be used in screening or for the treatment
CC of renal or urogenital cancer, e.g. bladder, cervical or prostate cancer.
CC Agents that may modulate this antigen could be used in the manufacture of
CC a medicament for the treatment of a tumour cell.
SQ Sequence 216 AA:

Query Match 37.9%; Score 1276; DB 1; Length 216;
Best Local Similarity 73.1%; Pred. NO. 1,44e-122;
Matches 158; Conservative 26; Mismatches 32; Indels 0; Gaps 0;

Db 1 TCNEGQQLGELNYRCDDTGTNDIPICVYKCLPTAPENKRIYSSANBPDRHFQ 60
QY 95 TCDEGQQLGELIDYRCDDAGTNDIPICVYKCLPTALENGRIYSGAEPDQYFFQ 154
Db 61 AVRFVNSGKIEGDEEMHCSDDGFWEKPKVCVEISCKSPVDYINGSPISOKIYKENER 120
QY 155 VYRFECNSGFKIEGQKEMHCSENGLMSNEKPCQVEISCLPPRYENDGDIYKRYKENER 214
Db 121 FOYKCMAGYESRGDAVCTESGWRPLPSCSEKSCNPITPNQDYSPLRIKHRTGDEITY 180
QY 215 FOYKCKGQFVYKERGDVATCGSGWNPQSCHEMTCLTFPIPNQIYTHRIKHRIIDDEIRY 274
Db 181 QCRNGFYPATRGNAKCTSTGWIAPRCLTKPCDYP 216
QY 275 ECKNGFYPATRSVSKCTITGWIAPRCLTKPCDYP 310

```

```

PR 06-OCT-1997; JP-272837.
PA (PROT-) PROTEGENE INC.
PA (SAGA) SAGAMI CHEM RES CENT.
PI Kato S, Sekine S;
DR WPI: 99-264019/22.
DR N-PSDB: X34737.
PT Human proteins with secretory signal sequences and nucleotide
PT sequences, useful in control of proliferation and differentiation of
PT cells
PS Claims 1; Page 55-58; 71pp; English.
CC This DNA encodes a protein having homology to human complement factor H,
CC which plays a role in the immunological mechanism involving the
CC complement reaction. The protein can also be used as an antigen for
CC preparing antibodies against the protein. The cDNA can be used as a probe
CC for gene diagnosis and the gene for gene therapy, as well as for large-
CC scale expression of the protein. The protein may also have immune
CC stimulating or suppressing activity, haematopoiesis regulating activity,
CC tissue growth activity, activin/inhibin activity, anti-inflammatory
CC activity, tumour inhibition activity, chemotactic/chemokinetic activity,
CC receptor/ligand activity, etc. The protein is identified by the presence
CC of a hydrophobic N-terminal secretory signal region. Knowledge of the
CC protein function is not required, as in e.g. methods of expression
SQ Sequence 578 AA:

```

```

Query Match 21.1%; Score 712; DB 1; Length 578;
Best Local Similarity 35.1%; Pred. NO. 1,36e-61;
Matches 124; Conservative 72; Mismatches 127; Indels 30; Gaps 25;

```

```

Db 54 YTCDEGQYLL-GRI-DYRECDAAGTNDIPICVYKCL-PTYLENGRIYSGAEPDQY 150
QY 94 YTCDEGQYLL-GRI-DYRECDAAGTNDIPICVYKCL-PTYLENGRIYSGAEPDQY 150
Db 108 ILMEEFOYKPKRYATADNGSSSITCLONG-WSTQ-PLCIKF-CMPVFNRSRKS-NG 163
QY 151 YFQYVYRFECNSGF-KIEQKE--MHCSSEGLMSNKPQVEISCLPPRYENDGDIYKLP 207
Db 164 MMEKLDITLDYECYDESSYGTNTDSIYCGEDGWSHLPTCVNSSSCGPPISNGDTT 223
QY 208 V-YKENERFOYKCKGQF--YKERGDA-VCTGSGWNPQSC-FEM-TC-LTYIPNGIYT 260
Db 224 SFPOKYILPMSRYEYCOQ--YELQSKYVTCNSGDSPPRCISMKPCEPFIQGHLY 282
QY 261 PHRIKHRIID-DEIRYECCKNGFPATRSVSKCTITGWIAPRCLTKPCDYPFQKHGRLY 318
Db 283 YENTRPYPVATGQSYTYCDONFVTPSGSYWDYIHCQDQWLPVYRCLRTCSKSDIET 342
QY 319 YVESRPYIPVIGKYEYTYCNGFTTTPSQSYWDYLRCTVNGWEPEVPLRQCI-FHY-V 376
Db 343 ENGFISSSSITILNKEIYKCKPGYATADNGSSSITCLONGWSAOPICIKF 395
QY 377 EYGESSYWMQRRYIEGQSAKVQCHSGYSLPNQDT--YYCTENGWSPPPRCVNI 427

```

```

RESULT 4
ID R13490 standard; Protein; 581 AA.
AC R13490;
DT 30-OCT-1991 (first entry)
DE Human C4 binding protein.
KW C4bp; monomer; complement protein; p10D.C4bp.3; SCR;
KW short consensus repeat.
OS Homo sapiens.
FH Key 1.32
FH peptide Location/Qualifiers
FH 1..32 /label=signal_peptide
FH 33..581 /label=C4bp
FH 33..93 /label=SCR8
FH 94..155 /label=SCR7
FH 156..219 /label=SCR6
FH region
FH region
FH region

```

```

FT region 220..279
FT /label= SCRS
FT region 280..345
FT /label= SCRA
FT region 346..406
FT /label= SCRA
FT region 407..464
FT /label= SCRA
FT region 465..523
FT /label= SCRA
FT domain 524..581
FT /label= C4bp-core
FT /note= "responsible for multimer assembly"
FT disulfide_bond 34..80
FT /note= "intradomain"
FT disulfide_bond 65..92
FT /note= "intradomain"
FT WO911461-A.
FT 08-AUG-1991.
FT 28-JAN-1991: U00567.
FT 26-JAN-1990: US-470888.
FT (BIOG-) BIGEN INC.
FT Pasek MP, Winkler G, Liu TR:
FT MPI: 91-252613/34.
FT DR N-PSDB: Q13242.
FT New C4 binding protein fusion proteins and DNA encoding them -
FT comprise assemblies of C4bp monomers linked to functional moiety,
FT e.g. Azr, useful as delivery vehicles in diagnosis and therapy
FT Example 1; Fig 1; 105pp; English.
FT This sequence was deduced from human hepatocyte (Hep G2) cDNA
FT obtained following PCR amplification. The protein is a monomer
FT containing 8 SCRs. Each SCR forms a looped domain due to the
FT presence of two intradomain disulfide bonds (only the disulfide
FT bonds of SCR8 are labelled in the Features Table). Within each SCR,
FT the first cysteine residue bonds with the third and the second
FT cysteine residue bonds with the fourth. This secondary structure is
FT responsible for the conformational flexibility of the C4bp monomer.
FT The invention covers fusion proteins in which the monomer sequence,
FT or a specified subfragment of it, e.g. having 5, 4, 3 or 1 SCR(s)
FT is fused to the C-terminal of a protein such as a viral receptor,
FT cell ligand, a bacterial, viral or parasitic immunogen, enzyme,
FT cytokine, toxin, etc. See also Q13243-51.
FT Sequence 581 AA:
SQ
Query Match 15.9%; Score 537; DB 1; Length 581;
Best Local Similarity 30.7%; Pred. No. 5,15e-43;
Matches 138; Conservative 85; Mismatches 182; Indels 45; Gaps 40:
Db 33 NC-GPPTLSFAPAMDITLRETRFKTGLTKYTCLPGVNHSHTQTLTNSDGEWV-YNT 90
:| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
2 DCKGPPPREN-SEILSGSWSEQLYSEGTQATYKCRPGY-RTLGTIVYCK-NGEMVSPNP 58
:| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
91 FCYI-KR-CRHPELRNG--OVELKIDLSFGSOIEFSGSEFFLTGSTSCQODRGV 146
:| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
59 SRICKRPPCGHGTGPGESEFLAVGSEPHGAKAVYTCDDESYQLGELDR-EC-DAD-G 115
:| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
147 WSHPELPOCELYCKPPDIRNGR-HSG--E-ENFYAYGFVSYSCDPRFSLGHASISCT 202
:| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
116 WTNDIPICEVYKCLPVTELENGRIVSGAEPDQYFGQVAFRCNSGFKIEGQKEMHS 175
:| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
203 VENETIGWRSPPTCEKITCRKPDVSHGEMVSGFPIYNTKDTIVRCQKGFVLKRGSSV 262
:| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
176 -EN--GLMSNEKPOCYEISCLPVRVNGDGY-LKPYKENERFOYCKGQGFYKKEGD 230
:| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
263 IHCADSKMNPSPACPEINCLPDIPHASMETYPRPTKEDVYVGVLYRCHPGKYP 322
:| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
231 AVC-TGSGWNP-QPSCCEMTCLT-PIYPNGIT-I-PRIKIRI--DDGI-RYECKNGYTP 282
:| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
323 TTDEPTIVICOKNLRTWTPYOGCEALCCPEKLNNGEIT-QH-RKS-RPANHC-VY-FYGD 377
:| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
283 ATRSPVS-KCIT-T-GWIPARCSLKPDPQFKHGRLYEESRRPYFVPIGKEYSYCD 340
:| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 378 EISCHETSFSAL-CGGDGTWSPRTSCGDICNFP-PKIAHGYKSSSSSFKEELI 435
:| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :

```

```

OY 341 N-GFTTSPQSQSYWDYINCTYWG--WEPEVP-CLRQCIFHYVEYSESSYWR-R-YTEGSAKV 396
:| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 436 YECDDKGYILV-GQAKLCSYSHWSAPAPQC 464
:| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
OY 397 -QCHSGTSLPNGODITYCTENGNSPP-PKC 424
:| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
RESULT 5
ID R28557 standard; peptide; 543 AA.
AC R28557:
DT 19-MAR-1993 (first entry)
DE CRI-4 (99H, 103E) analogue.
KW short consensus repeat; regulator of complement activation;
KW C3b binding; C4b binding; human complement type 1 receptor.
OS Homo sapiens.
FH Key Location/Qualifiers
FT region 1..60
FT /label= SCR-1
FT region 61..122
FT /label= SCR-2
FT region 451..510
FT /label= SCR-8
FT region 511..543
FT /label= SCR-9
FT /note= "TRUNCATED"
FT misc_difference 99
FT /note= "Ser substituted by His (SCR-9)"
FT misc_difference 103
FT /note= "Thr substituted by Glu (SCR-9)"
FT EP-512733-A.
FT PD 11-NOV-1992.
FT 28-APR-1992: 303826.
FT 03-MAY-1991: US-695514.
FT (UNIV ) UNIV WASHINGTON.
FT PI Atkinson JP, Hourcade D, Kiyoh M;
FT MPI: 92-375009/46.
FT DR Complement actively regulator protein analogues - useful for
FT treating auto-immune diseases, to suppress transplant rejection,
FT for diagnosis etc.
FT Claim 11: Fig 2 and R11810: 23pp; English.
FT The cDNA clone designated CRI-4 was described in J Exp Med. (1988)
FT 168:1255-1270. It encodes the first 8 and a half amino terminal
FT SCRs of CRI. The invention concerns analogues of "regulator of
FT complement activation" proteins or truncated, hybrid or recombinant
FT forms of them. CRI-4 is a preferred truncated form and a number of
FT specified substitution variants of it are claimed in which certain
FT positions in SCR-2 which have been identified as important for the
FT degree of C3b- and C4b-binding are substituted by amino acids from
FT the corresponding positions in SCR-9. The specification does not
FT contain the CRI-4 sequence; the sequence given here was constructed
FT from the full-length CRI amino acid sequence having GEMSEQ
FT accession number R11810 and descriptions in the disclosure.
FT Sequence 543 AA:
SQ
Query Match 13.2%; Score 445; DB 1; Length 543;
Best Local Similarity 32.5%; Pred. No. 2,21e-33;
Matches 105; Conservative 72; Mismatches 106; Indels 40; Gaps 34:
Db 22 PPIGTLYNEGRPGYSGR-PSIICLNKSNVWTGAK-DR-CRRKSCRNPDP-PVNGM-VHV 76
:| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
OY 24 YSEGTQATYKCRPGYRILGTIVYKC-KNGEMVSPNPSCRKRCGHPGDPFSSFLAV 82
:| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 77 IKGIQFSGQIKYSCTKGKYLIGHSS-AECIISGDTIWMNETPICDRIPGLPPT-ITNG 134
:| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
OY 83 GSEFEFGAKVYVYTDGEGYQLGELDR-EC-DADG--WTNDIPICEVYKC-LPVTELENG 137
:| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 135 DFIS--INRE-NHYGSAVYTRCNPGSGGKRVFLVGPSTLYCSNDQVQVINGSPAPQC 191
:| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
OY 138 RIVSGAEPDDEYFGQVAFRCN--SGF-KI-E-GQKEMHC-S-EN--GLMSNEKPOC 187
:| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 192 IIPNKCTPVENGLIVSDNLSFLSNEVVEFRCPQPVVMGPRRVYCOALNKKPELPS 251
:| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :

```


AC R28553;
 DT 19-MAR-1993 (first entry)
 DE CRI-4 (85R, 87N) analogue.
 KW short consensus repeat; regulator of complement activation;
 OS Homo sapiens.
 FH C3b binding; C4b binding; human complement type 1 receptor.
 FT Key
 FT region Location/Qualifiers
 FT 1..60
 FT /label= SCR-1
 FT 61..122
 FT /label= SCR-2
 FT 451..510
 FT /label= SCR-8
 FT 511..543
 FT /label= SCR-9
 FT /note= "TRUNCATED"
 FT misc_difference 85
 FT /note= "Gln substituted by Arg (SCR-9)"
 FT misc_difference 87
 FT /note= "Lys substituted by Asn (SCR-9)"
 PN EP-512733-A.
 PD 11-NOV-1992.
 PF 28-APR-1992; 303826.
 PR 03-MAY-1991; US-695514.
 PA (UNIW) UNIV WASHINGTON.
 PI Atkinson JP, Hourcade D, Krych M;
 DR WPI: 92-375009/46.
 PT Complement activity regulator protein analogues - useful for
 PT treating auto-immune diseases, to suppress transplant rejection,
 PT for diagnosis etc.
 PS Claim 11; Fig 2 and R11810; 23pp; English.
 CC The cDNA clone designated CRI-4 was described in J.Exp.Med.(1988)
 CC 168:1255-1270. It encodes the first 8 and a half amino terminal
 CC SCRs of CRI. The invention concerns analogues of "regulator of
 CC complement activation" proteins or truncated, hybrid or recombinant
 CC forms of them. CRI-4 is a preferred truncated form and a number of
 CC specified substitution variants of it are claimed in which certain
 CC positions in SCR-2 which have been identified as important for the
 CC degree of C3b- and C4b-binding are substituted by amino acids from
 CC the corresponding positions in SCR-9. The specification does not
 CC contain the CRI-4 sequence; the sequence given here was constructed
 CC from the full-length CRI amino acid sequence having GENSESQ
 CC accession number R11810 and descriptions in the disclosure.
 SQ Sequence 543 AA;

Query Match 13.1%; Score 441; DB 1; Length 543;
 Best Local Similarity 32.0%; Pred. No. 5.77e-33;

Matches 103; Conservative 73; Mismatches 108; Indels 38; Gaps 33;

DB 22 PPIGTLYNTECRPGYSGR-PFSITCLKNSVWTGAK-DR-CRRKSCNPPD-PVNGM-VHY 76
 QY 24 YSEGTATYRCRPGYRFLGTIVKVC-KNGEWPVNSNRICKRKPCHPDPTFGSRLAV 82
 DB 77 IKGIQFSGRINYSCTGYRLIGSSSATCIISGTYVMSNETPICDRIPGGLPPT-ITNGD 135
 QY 83 GSEFEFGAKVYTCDSYQLLGEIDYR-ECDAAG--WTNDIPCEVVKC-LPVTELENGR 138
 DB 136 FIS--TNRE-NHYGSAVYTRCNPSSGGRKVFELVGEPSITCSDNDQVGTSGAPAOI 192
 QY 139 IVSGAEPDDEYFGQVYFEKCN--SGF-KI-E--GQKEMHC-S-EN--GLMSNEKPCOV 188
 DB 193 IPKRCPPVNGENILVSDNRSLFSLNEVEVEFRQPYFVVKGPRRYKCOALNKEPELPS 252
 QY 189 EIS-CLPPRVENDGITY-LKPYKENERFOYKCKOGFYVKEKGDVAVCG-SGMNQ-PSC 244
 DB 253 SRV-CQPPPVLAERTQDKDNFSPGOEVFYSCEPG-Y-DLRGAASMRCTPOGWSPPAA 309
 QY 245 EEMTCLTP-YIINGITTPH-RIKHRIDDIRCKNGEYFATRSVPS-KCITTG-WIDA- 299
 DB 310 PTCEVKSQCDDEMGQLNGRVLF 331
 QY 300 PRCSLKPCD-F-PQFKHGRLYY 319

RESULT 11
 ID R28560; standard; peptide; 543 AA.
 AC R28560;
 DT 19-MAR-1993 (first entry)
 DE CRI-4 (114S) analogue.
 KW short consensus repeat; regulator of complement activation;
 OS Homo sapiens.
 FH C3b binding; C4b binding; human complement type 1 receptor.
 FT Key
 FT region Location/Qualifiers
 FT 1..60
 FT /label= SCR-1
 FT 61..122
 FT /label= SCR-2
 FT 451..510
 FT /label= SCR-8
 FT 511..543
 FT /label= SCR-9
 FT /note= "TRUNCATED"
 FT misc_difference 114
 FT /note= "Asp substituted by Ser (SCR-9)"
 PN EP-512733-A.
 PD 11-NOV-1992.
 PF 28-APR-1992; 303826.
 PR 03-MAY-1991; US-695514.
 PA (UNIW) UNIV WASHINGTON.
 PI Atkinson JP, Hourcade D, Krych M;
 DR WPI: 92-375009/46.
 PT Complement activity regulator protein analogues - useful for
 PT treating auto-immune diseases, to suppress transplant rejection,
 PT for diagnosis etc.
 PS Claim 11; Fig 2 and R11810; 23pp; English.
 CC The cDNA clone designated CRI-4 was described in J.Exp.Med.(1988)
 CC 168:1255-1270. It encodes the first 8 and a half amino terminal
 CC SCRs of CRI. The invention concerns analogues of "regulator of
 CC complement activation" proteins or truncated, hybrid or recombinant
 CC forms of them. CRI-4 is a preferred truncated form and a number of
 CC specified substitution variants of it are claimed in which certain
 CC positions in SCR-2 which have been identified as important for the
 CC degree of C3b- and C4b-binding are substituted by amino acids from
 CC the corresponding positions in SCR-9. The specification does not
 CC contain the CRI-4 sequence; the sequence given here was constructed
 CC from the full-length CRI amino acid sequence having GENSESQ
 CC accession number R11810 and descriptions in the disclosure.
 SQ Sequence 543 AA;

Query Match 13.0%; Score 439; DB 1; Length 543;
 Best Local Similarity 32.0%; Pred. No. 9.31e-33;

Matches 103; Conservative 73; Mismatches 108; Indels 38; Gaps 33;

DB 22 PPIGTLYNTECRPGYSGR-PFSITCLKNSVWTGAK-DR-CRRKSCNPPD-PVNGM-VHY 76
 QY 24 YSEGTATYRCRPGYRFLGTIVKVC-KNGEWPVNSNRICKRKPCHPDPTFGSRLAV 82
 DB 77 IKGIQFSGRINYSCTGYRLIGSSSATCIISGTYVMSNETPICDRIPGGLPPT-ITNGD 135
 QY 83 GSEFEFGAKVYTCDSYQLLGEIDYR-ECDAAG--WTNDIPCEVVKC-LPVTELENGR 138
 DB 136 FIS--TNRE-NHYGSAVYTRCNPSSGGRKVFELVGEPSITCSDNDQVGTSGAPAOI 192
 QY 139 IVSGAEPDDEYFGQVYFEKCN--SGF-KI-E--GQKEMHC-S-EN--GLMSNEKPCOV 188
 DB 193 IPKRCPPVNGENILVSDNRSLFSLNEVEVEFRQPYFVVKGPRRYKCOALNKEPELPS 252
 QY 189 EIS-CLPPRVENDGITY-LKPYKENERFOYKCKOGFYVKEKGDVAVCG-SGMNQ-PSC 244
 DB 253 SRV-CQPPPVLAERTQDKDNFSPGOEVFYSCEPG-Y-DLRGAASMRCTPOGWSPPAA 309
 QY 245 EEMTCLTP-YIINGITTPH-RIKHRIDDIRCKNGEYFATRSVPS-KCITTG-WIDA- 299
 DB 310 PTCEVKSQCDDEMGQLNGRVLF 331
 QY 300 PRCSLKPCD-F-PQFKHGRLYY 319

RESULT 12
ID R28558 standard; peptide: 543 AA.
AC R28558;
DT 19-MAR-1993 (first entry)
DE CRI-4 (109N, 110A, 111A, 112H) analogue.
KW Short consensus repeat; regulator of complement activation;
C3b binding; C4b binding; human complement type 1 receptor.
OS Homo sapiens.
FH Key Location/Qualifiers
FT region 1..60
FT /label= SCR-1
FT region 61..122
FT /label= SCR-2
FT region 451..510
FT /label= SCR-8
FT region 511..543
FT /label= SCR-9
FT /note= "TRUNCATED"
FT misc_difference 109
FT /note= "Asp substituted by Asn (SCR-9)"
FT misc_difference 110
FT /note= "Thr substituted by Ala (SCR-9)"
FT misc_difference 111
FT /note= "Val substituted by Ala (SCR-9)"
FT misc_difference 112
FT /note= "Ile substituted by His (SCR-9)"
FT EP-512733-A.
PN 11-NOV-1992.
PD 28-APR-1992; 303826.
PR 03-MAY-1991; US-695514.
PA (UNITV) UNIV WASHINGTON.
PI Atkinson JP, Hourcade D, Krych M;
DR WPI: 92-375009/46.
PT Complement actively regulator protein analogues - useful for
treating autoimmune diseases, to suppress transplant rejection,
PT for diagnosis etc.
PS Claim 11; Fig 2 and R1810; 23pp; English.
CC The cDNA clone designated CRI-4 was described in J. Exp. Med. (1988)
CC 168:1235-1270. It encodes the first 8 and a half amino terminal
CC SCRs of CRI. The invention concerns analogues of "regulator of
CC complement activation" proteins or truncated, hybrid or recombinant
CC forms of them. CRI-4 is a preferred truncated form and a number of
CC specified substitution variants of it are claimed in which certain
CC positions in SCR-2 which have been identified as important for the
CC degree of C3b- and C4b-binding are substituted by amino acids from
CC the corresponding positions in SCR-9. The specification does not
CC contain the CRI-4 sequence; the sequence given here was constructed
CC from the full-length CRI amino acid sequence having GENESEQ
CC accession number R1810 and descriptions in the disclosure.
CC Sequence 543 AA:
SQ

Query Match 13.0%; Score 439; DB 1; Length 543;
Best Local Similarity 32.0%; Pred. No. 9.31e-33;
Matches 103; Conservative 72; Mismatches 109; Indels 38; Gaps 33;

DB 253 SRV-CQPPDYLAERTQDRDNFSGOEVYSCPEG-Y-DLRGAASRCTPGQWSPPAA 309
QY 245 EEMTCLTP-YIPNGIYTPH-RIKHRIDDEIRYECCKNGFPATRSPPVS-KCTITG-WIPA- 299
DB 310 PTCEVKSDDDEMGQLNGRYLF 331
QY 300 PRCSLKPCD-F-POFKHGRILY 319

RESULT 13
ID R28571 standard; peptide: 543 AA.
AC R28571;
DT 19-MAR-1993 (first entry)
DE CRI-4 (364-367 NAAH) analogue.
KW Short consensus repeat; regulator of complement activation;
C3b binding; C4b binding; human complement type 1 receptor.
OS Homo sapiens.
FH Key Location/Qualifiers
FT region 1..60
FT /label= SCR-1
FT region 61..122
FT /label= SCR-2
FT region 451..510
FT /label= SCR-8
FT region 511..543
FT /label= SCR-9
FT /note= "TRUNCATED"
FT misc_difference 364..367
FT /note= "Met substituted with NAAH from
SCR-8-9 to increase C3b binding"
FT EP-512733-A.
PN 11-NOV-1992.
PD 28-APR-1992; 303826.
PR 03-MAY-1991; US-695514.
PA (UNITV) UNIV WASHINGTON.
PI Atkinson JP, Hourcade D, Krych M;
DR WPI: 92-375009/46.
PT Complement actively regulator protein analogues - useful for
treating autoimmune diseases, to suppress transplant rejection,
PT for diagnosis etc.
PS Example 8; Page 18 and R1810; 23pp; English.
CC The cDNA clone designated CRI-4 was described in J. Exp. Med. (1988)
CC 168:1235-1270. It encodes the first 8 and a half amino terminal
CC SCRs of CRI. The invention concerns analogues of "regulator of
CC complement activation" proteins or truncated, hybrid or recombinant
CC forms of them. CRI-4 is a preferred truncated form and a number of
CC specified substitution variants of it are disclosed in which certain
CC positions in SCR-5-6 are substituted by amino acids from
CC the corresponding positions in SCRs which are involved in C3b- and
CC C4b-binding. The substitution variant given here has increased C3b-
CC binding. The specification does not contain the CRI-4 sequence;
CC the sequence given here was constructed from the full-length CRI
CC amino acid sequence having GENESEQ accession number R1810 and
CC descriptions in the disclosure.
CC Sequence 543 AA:
SQ

Query Match 13.0%; Score 438; DB 1; Length 543;
Best Local Similarity 31.9%; Pred. No. 1.18e-32;
Matches 119; Conservative 79; Mismatches 127; Indels 48; Gaps 39;

QY 189 EIS-CLPFRVENGDDIY-LKPVYKENERFOYCKOGFYFKRGDAVCTG-SGMNPQ-PSC 244
 DB 253 SRV-COPPPDLHAERTORDKDNFSPGOEYVYSCPEG-Y-DLRGAASMRCTPOGDSPPA 309
 QY 245 EEMTCLTP-YIPNGIYTPH-RIKHRIDELRYECKNGFYPATRSPVS-KCTITG-WIPA- 299
 DB 310 PTCEVKSCLD-DEM-GOLL-NGR-VLFPVNLQGLARVDFVCEGFOGKGSASV-CYLA 362
 QY 300 PRCSLKPCDDFOFKHGRLYEESRPFYFV-PIGKRYSYCCDNGF--TTPSOSYMDYLR 355
 DB 363 GNAHMSVYVC 375
 QY 356 CTYNGMEPEVP-C 367

RESULT 14
 ID R28550 standard; peptide; 543 AA.
 AC R28550:
 DT 19-MAR-1993 (first entry)
 DE CRI-4 (64K) analogue.
 KW short consensus repeat; regulator of complement activation;
 KW C3b binding; C4b binding; human complement type 1 receptor.
 OS Homo sapiens.
 FH Key
 FT region 1..60
 FT /label= SCR-1
 FT region 61..122
 FT /label= SCR-2
 FT region 451..510
 FT /label= SCR-8
 FT region 511..543
 FT /label= SCR-9
 FT /note= "TRUNCATED"
 FT misc_difference 64
 FT /note= "Arg substituted by Lys (SCR-9)"
 PN EP-512733-A.
 PF 11-NOV-1992.
 PF 28-APR-1992: 303826.
 PR 03-MAY-1991: US-695514.
 PA (UNIW) UNIV WASHINGTON.
 PI Atkinson JP, Hourcade D, Krych M;
 WP1: 92-375009/46.
 PT Complement activator/regulator protein analogues - useful for
 PT treating autoimmune diseases, to suppress transplant rejection,
 PT for diagnosis etc.
 PS Claim 11: Fig 2 and R11810; 23pp; English.
 CC The cDNA clone designated CRI-4 was described in J.Exp.Med.(1988)
 CC 168:1255-1270. It encodes the first 8 and a half amino terminal
 CC SCRs of CRI. The invention concerns analogues of "regulator of
 CC complement activation" proteins or truncated, hybrid or recombinant
 CC forms of them. CRI-4 is a preferred truncated form and a number of
 CC specified substitution variants of it are claimed in which certain
 CC positions in SCR-2 which have been identified as important for the
 CC degree of C3b- and C4b-binding are substituted by amino acids from
 CC the corresponding positions in SCR-9. The specification does not
 CC contain the CRI-4 sequence; the sequence given here was constructed
 CC from the full-length CRI amino acid sequence having GENSSEQ
 CC accession number R11810 and descriptions in the disclosure.
 SQ Sequence 543 AA;

Query Match 13.0%; Score 437; DB 1; Length 543;
 Best Local Similarity 32.0%; Pred. No. 1.50e-32;
 Matches 103; Conservative 72; Mismatches 109; Indels 38; Gaps 33;

DB 22 PPIGTLYNECRPGYSGR-PSITICLDNSVWTGAK-DR-CRRKSKRPPD-PVNGM-VHV 76
 QY 24 YSEGTQATYKCRPGYRRLGTIVKVC-KNGEWPSPSRICKRRCGHPGTGPGSFLAV 82
 DB 77 IKGIQFSSQIKYSCGKRYLIGSSSATCIISGDPVINDNEPTICDRIPGGLPPT-ITND 135
 QY 83 GSEFEFAKVVYTCDEGYQLIGEDYR-ECDADG-WTNDIPICEVVC-LPVTLENGR 138

DB 136 FIS--TNRE-NHYGSVYTYRCNPGSGGRKYFELVGEPSYCTSDNDQYIGSGAPOCI 192
 QY 139 IYSAAEAPDQYFYFQVVRFEEN--SGF-KI-E--GQEMHC-S-EN--GLMSNEKQCV 188
 DB 193 INKCTPRVENVGLIVSDNRSLFSLNEVEYRCOPVFMKGPBRVYKQALNKKMBELPSC 252
 QY 189 EIS-CLPFRVENGDDIY-LKPVYKENERFOYCKOGFYFKRGDAVCTG-SGMNPQ-PSC 244
 DB 253 SRV-COPPPDLHAERTORDKDNFSPGOEYVYSCPEG-Y-DLRGAASMRCTPOGDSPPA 309
 QY 245 EEMTCLTP-YIPNGIYTPH-RIKHRIDELRYECKNGFYPATRSPVS-KCTITG-WIPA- 299
 DB 310 PTCEVKSCLDDEM-GOLLNGRVL 331
 QY 300 PRCSLKPCDD-F-PQKHGRLY 319

RESULT 15
 ID R11982 standard; Protein; 1537 AA.
 AC R11982:
 DT 25-JUN-1991 (first entry)
 DE Partial human complement type 1 receptor.
 KW complement system; C3b/C4b receptor; CRI; allergic reaction;
 KW immune response; long homologous repeat; LHR.
 OS Homo sapiens.
 FH Key
 FT region 1..438
 FT /label= LHR-B
 FT region 439..891
 FT /label= LHR-C
 FT region 892..1341
 FT /label= LHR-D
 FT region 1495..1498
 FT /note= "positively-charged; preceded by hydrophobic
 FT sequence."
 FT 1521..1526
 FT /note= "has 67 per cent homology to site of protein
 FT kinase C phosphorylation in the EGF
 FT receptor"

PN W09105047-A.
 PD 18-APR-1991.
 PF 25-SEP-1990: 005454.
 PR 26-SEP-1989: US-412745.
 PR 26-SEP-1990: US-912349.
 PA (TCEL-) T CEL SCI INC.
 PA (UYJO) JOHNS HOPKINS UNIVERSITY.
 PA (BRIG-) BRIGHAM AND WOMEN'S HOSPITAL.
 PI Fearon DT, Klickstein LB, Wong WW, Carson GR, Hoh M, Concino MF,
 PI Makrides SC, Marsh HC;
 DR WP1: 91-132854/18.
 DR N-PSDB: Q11643.
 PT Human complement receptor type 1 gene, encoded proteins and
 PT fragments - for treatment of immune disorders, myocardial infarct,
 PT damage due to inflammation and in treatment of thrombosis
 PS Disclosure; Fig 5; 234pp; English.
 CC This sequence comprises three of the four tandem, direct, long
 CC homologous repeats of the full-length F allozyme of CRI. LHR-A is
 CC absent. Each LHR might represent a single C3b/C4b binding domain,
 CC making the receptor multivalent. The LHR's are composed of 7 short
 CC consensus repeats of 60-70 residues resembling the SCR's of other
 CC C3/C4 binding proteins. The protein and fragments of it having C3b
 CC and/or C4b binding activity can be used to treat immune disorders
 CC or disorders involving inappropriate complement activity.
 CC See also Q11642.
 SQ Sequence 1537 AA;

Query Match 13.0%; Score 439; DB 1; Length 1537;
 Best Local Similarity 32.6%; Pred. No. 9.31e-33;
 Matches 105; Conservative 73; Mismatches 106; Indels 38; Gaps 33;

DB 461 PPIGTLYNECRPGYSGR-PSITICLD-NLYWSSPKDYCKRCKSKRPPD-PVNGM-VHV 516
 QY 24 YSEGTQATYKCRPGYRRLGTIVKVC-KNGEWPSPSRICKRRCGHPGTGPGSFLAV 83

```

Db 517 TDIOVGRINSCITGHLIGHSS-ABCIISGNTAHMSTKRPICQRIPOGLPPT-IANGD 574
QY 84 SEFEFGAKVYVTCDEGYOLLEIDYREC-DAD-G-WTNDIPICEVVKC-LPVTLENGR 138
Db 575 FTS--TNRE-NFHYGSVYTYRCNLGSRGRKRYFELVGEPSIYCTSNDDQVGIMSGPAPQCI 631
QY 139 IYSGAEPDQEVYFGQVYRFECH--S-GFKI-E-GQKEMHC-S-EN--GLMSNEXKPCV 188
Db 632 IPNKCTPPNVENGIIIVSDNLSLFSINEVEFRPCOPGYVMKGRPRYKCOALNKMEBELPSC 691
QY 189 EIS-CLPPRENGDSIX-LKPYKENERFOYKCKOGFYKKEGDAVCTG-SGWNPO-PSC 244
Db 692 SRV-QOPPELHGHMTPSHDNFSPGOEVFYSCEPG-Y-DLRGAASLHCTPRGDMSPEA 748
QY 245 EEMTCLTP-YIPNGIYTP-HRIKHRIDDEIRYECKNGFYPATRSPVS-KCITIIG-WIP-A 299
Db 749 PRCAYKSCDDFLGOLPHGRVLF 770
QY 300 PRCSLKPCD-F-PQFKHGRLY 319

```

Search completed: Thu Jun 8 21:49:27 2000
 Job time : 17 secs.

Release 3.1A John F. Collins, Biocomputing Research Unit
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

```

Mpsrch_pp  protein - protein database search, using Smith-Waterman algorithm
Run on:      Thu Jun  8 21:49:45 2000;      MasPar time 24.14 Seconds
Tabular output not generated.      836.361 Million cell updates/sec

```

```

Title: >US-09-316-163-14
Description: (1-428) from US09316163.pep
Perfect Score: 3371
Sequence: 1 EDCKGPPPRENSELSSWS.....DIYCTENGMSPPCKVRIK 428

```

Scoring table: PAM 150
Gap 11

Searched: 142080 seqs, 47172406 residues

Post-processing: Minimum Match 08
 Existing first 45 summaries

Database: plr62

Statistics: Mean 45.494; Variance 73.749; scale 0.617

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	2834	84.1	1234	1	NBMSH	complement factor H p	0.00e+00
2	2396	71.1	1231	1	NBHHH	complement factor H p	0.00e+00
3	2389	70.9	449	1	NBHHS	complement factor H p	0.00e+00
4	1720	51.0	669	2	S65551	factor H - bovine (fr	0.00e+00
5	1136	33.7	452	2	A35068	complement factor H-r	5.17e-25
6	1129	33.5	303	2	H35068	apolipoprotein H-rela	3.28e-24
7	790	23.4	808	2	D35069	complement factor H-r	1.40e-16
8	683	20.3	1053	2	S46199	probable complement r	1.28e-13
9	646	19.2	331	2	A45222	complement factor H-r	2.32e-12
10	564	16.7	537	1	S53711	C4Bp alpha chain prec	5.53e-10
11	537	15.9	597	1	NBHUC4	C4b-binding protein a	2.52e-99
12	494	14.7	558	2	S57953	C4Bp protein alpha ch	8.78e-89
13	474	14.1	610	1	I46001	C4b-binding protein a	6.57e-84
14	470	13.9	676	2	A45900	complement C3b recept	6.17e-83
15	447	13.3	343	2	G35070	apolipoprotein H-rela	2.31e-77
16	449	13.3	661	2	I36937	complement receptor 1	7.60e-78
17	449	13.3	2014	2	I36936	complement receptor 1	7.60e-78
18	446	13.2	2039	2	A28507	complement C3b/C4b re	4.04e-77
19	446	13.2	2489	2	I73012	complement C3b/C4b re	4.04e-77
20	435	12.9	327	2	A40455	complement factor H-r	1.82e-74
21	434	12.9	330	2	I56100	factor H homolog - hu	3.17e-74
22	434	12.9	330	2	I72653	factor H homolog - hu	3.17e-74
23	434	12.9	469	1	NBMSC4	C4b-binding protein a	3.17e-74

45	345	10.2	384	2	S01896	membrane cofactor pro	4.57e-55
44	345	10.2	377	2	I54479	membrane cofactor pro	4.57e-55
43	345	10.2	369	2	I54998	membrane cofactor pro	4.57e-55
42	348	10.3	668	2	A26606	endothelial leukocyte	9.07e-55
41	335	10.5	610	2	A42603	endothelial leukocyte	2.07e-55
40	357	10.6	618	2	B42755	E-selectin precursor	7.00e-55
39	359	10.6	612	2	S33046	endothelial leukocyte	2.37e-55
38	327	10.6	610	2	A50466	endothelial leukocyte	2.37e-55
37	364	10.8	345	1	AP0465	apolipoprotein H prec	1.57e-55
36	365	10.8	345	1	NHNU	apolipoprotein H prec	1.57e-55
35	363	10.8	263	1	C36838	complement control pr	2.71e-55
34	368	10.9	768	2	A42753	P-selectin precursor	1.79e-55
33	374	11.1	768	2	I33821	P-selectin - rat	6.84e-66
32	375	11.1	345	1	NEMS	apolipoprotein H prec	3.97e-66
31	390	11.6	830	2	A30359	P-selectin precursor	3.97e-66
30	395	11.7	263	1	WNVZSP	apolipoprotein H homo	4.07e-66
29	406	12.0	428	2	A24924	complement C3b/C4b re	2.89e-66
28	409	12.1	497	2	JC2050	complement regulatory	3.20e-68
27	411	12.2	579	2	A56740	sperm-eig recognition	1.06e-68
26	426	12.6	560	2	T16833	hypothetical protein	1.67e-72
25	431	12.8	270	2	I37278	complement factor H-r	1.67e-72
24	432	12.8	233	2	I37388	complement factor H-r	1.67e-72
23	436	12.6	243	2	I37388	complement factor H-r	1.67e-72

ALIGNMENTS

```

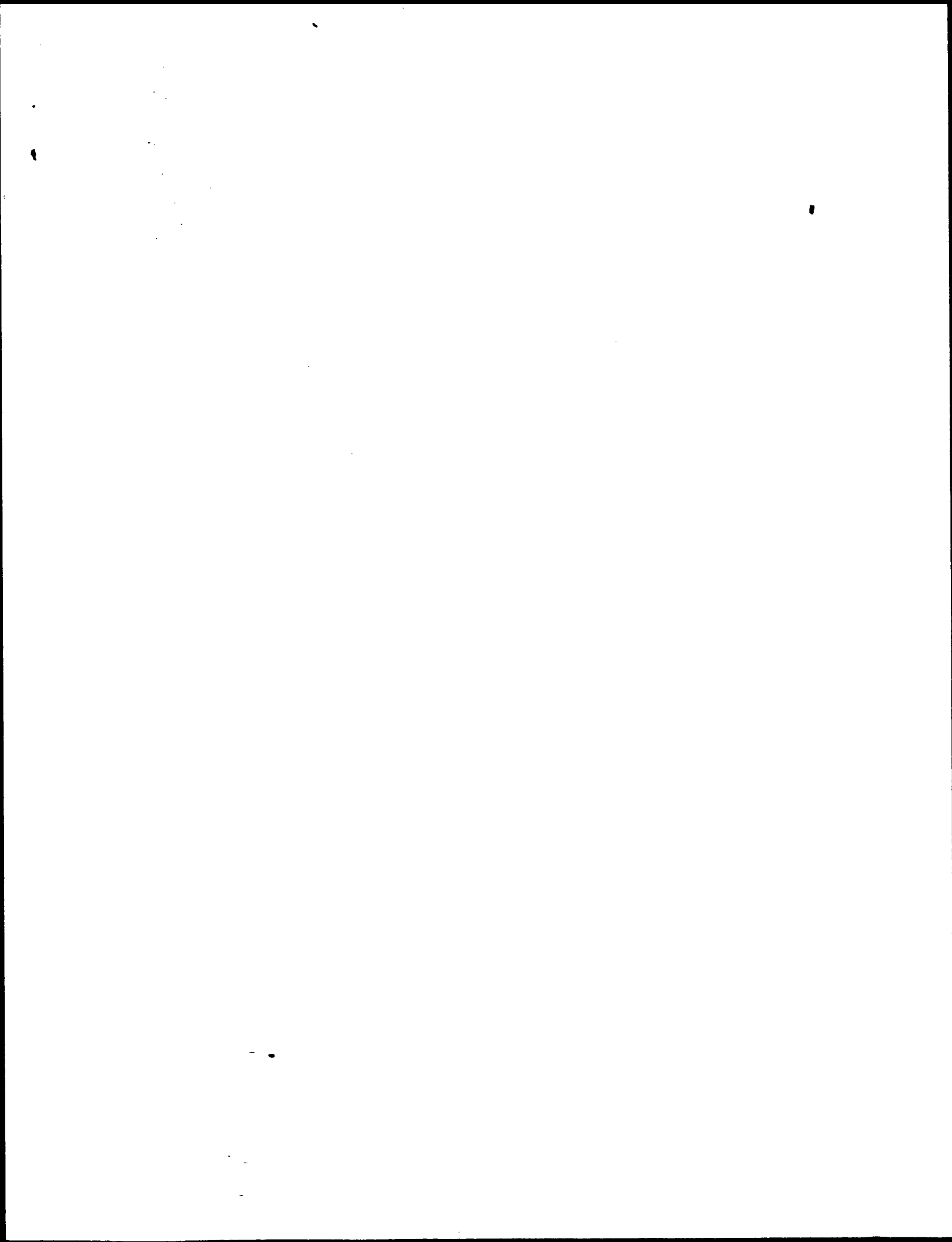
ENTRY      1
TITLE      NBKMSH
ALTERNATE_NAMES  #type complete
ORGANISM   complement factor H precursor - mouse
DATE       protein beta-1-H
           #formal_name Mus musculus #common_name house mouse
           30-Sep-1987 #sequence_revision 30-Sep-1987 #text-change
           22-Jun-1999
ACCESSIONS A26154; 149711; 149728
REFERENCE   A26154
#authors   Kristensen, T.; Tack, B.F.
#journal   Proc. Natl. Acad. Sci. U. S. A. (1986) 83:3963-3967
#title     Murine protein H is comprised of 20 repeating units, 61 amino
           acids in length.
#cross-references M01D:8623353
#accession  A26154
#molecule_type mRNA
#residues  1-1234 #label KRI
REFERENCE   #cross-references GB:M12660; NID:g193724; PIDN:AAA37759.1; PID:g387181
#authors   Natsuumi-Sakai, S.; Nonaka, M.; Nonaka, M.; Harada, Y.
#journal   J. Immunol. (1990) 144:358-362
#title     Demonstration of an unusual allelic variation of mouse factor
           H by the complete cDNA sequence of the H.2 allotype.
#cross-references M01D:90111033
#accession  I49711
#status     preliminary; translated from GB/EMBL/DBJ
#molecule_type mRNA
#residues  1-18 #label RES
#cross-references GB:M31979; NID:g193726; PIDN:AAA37762.1; PID:g193729
REFERENCE   I49728
#authors   Munoz-Canoves, P.; Tack, B.F.; Vlk, D.P.
#journal   Biochemistry (1989) 28:9891-9897
#title     Analysis of complement factor H mRNA expression:
           Dexamethasone and IFN-gamma increase the level of H in L
           cells.
#cross-references M01D:90148935
#accession  I49728
#status     preliminary; translated from GB/EMBL/DBJ
#molecule_type mRNA
#residues  1-19 #label RE2
#cross-references GB:J02891; NID:g193805; PIDN:AA37795.1; PID:g553926
COMMENT     TWO codominant alleles of factor H are present in mice.
COMMENT     Factor H functions as a cofactor in the inactivation of C3b by
COMMENT     serine proteinase I and also increases the rate of dissociation
COMMENT     of the C3bb complex (C3 convertase) and the (C3b)nb complex (C3
COMMENT     convertase) in the alternative complement pathway.
GENETICS

```


GENETICS	GENE	SB1	putative complement-regulatory plasma protein from barred sand bass (Paralichthys obsoletus).
CLASSIFICATION	g1ycoprotein	g1ycoprotein	
FEATURE			
89-145	#domain complement factor H repeat	#label FH01	
334-389	#domain complement factor H repeat	#label FH02	
450-502	#domain complement factor H repeat	#label FH03	
569-624	#domain complement factor H repeat	#label FH04	
743-802	#domain complement factor H repeat	#label FH05	
935-989	#domain complement factor H repeat	#label FH06	
993-1052	#domain complement factor H repeat	#label FH07	
SUMMARY	#length 1053	#molecular-weight 117597	#checksum 8482
Query Match	20.3%	Score 683;	DB 2; Length 1053;
Best Local Similarity	31.7%	Pred. No. 1,286-135;	
Matches 132;	Conservative 76;	Mismatches 180;	Indels 28; Caps 22;
Db	50	YPGGQVRRGCVNGV-S-GFRLKYCEGK-ETRGAK-CQPRSCGHPGDAQFADFLHAE	105
Qy	24	YSEGTQATYKCRPGYRTLTGIVKVCNKGWVPSPSRICRKRPGHGDTPFGSFLRAG	83
Db	106	NDFVSGKVVYICQKQYQWMSRINRYCAEAGDGVVPCESQGC-PLIHVDNNGVIG	164
Qy	84	SEFEGAKVYICDDESYQLGELIDYECADGNTNDIPICEVYKCLPYTELENGRIYSGA	143
Db	165	--PEAT-PGNVYRSCSRSEILDSPELYCDERDMSGVPKCAITCAIPIENGAV	221
Qy	144	APPDQYHYGVVYRFGCNGSGFKI-EGQKEMHCSENGLMSNEKPOQVEISCLPREVENDG	202
Db	222	PGAIRKKNVDLHYCDAPFKHIDPPSTICIKGIAEWSPTPLCESIKRLIMDTRY	201
Qy	203	ITLKPYEKNERQYCKGQFYKKEGDA-VCTG--SGMNPQSCCEMTLTPIYIPNGIT	259
Db	282	EPAYRNLFPGEFLKATYCATSWISIRPOEFSVYTTQDNGEMSLRPFCEVRCSNRPEH	341
Qy	260	TP-HRKHHIDDIRK-CKNGFYPAIR-SPV-SKCTITG-WLPAPKCSLAKPDDFPQFKH	314
Db	342	VDSWDVRSWERY--TLDNTRYKCRKRYKRTGVTWA--TCGRNGMPNPLCEVKTCSK	396
Qy	315	GRLYEESRRPYPPVPIGKEYSYCCNGFTTPSQSYVDYLRCTVNGWEPVPC-LRQCIF	373
Db	397	ENIQDAVYIGTKQITNLNOKALYACGEG-N--RGRITTLGCGENGSGDKRT-VK	448
Qy	374	HYVEYESSYWMRR-YIEGOSAKVOCHSGSLPNDQTYCTENGWSPSPKCAIRK	428
RESULT	9		
ENTRY	A45222	#type complete	
TITLE		complement factor H-related protein DOWN16 precursor - human	
ORGANISM		#formal_name Homo sapiens #common_name man	
DATE		30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 24-Sep-1998	
ACCESSIONS	A45222; S29609		
REFERENCE	A45222		
AUTHORS	Skerka, C.; Kuhn, S.; Gunther, K.; Lingelbach, K.; Zipfel, P.F.		
#journal	J. Biol. Chem. (1993) 268:2904-2908		
#title	A novel short consensus repeat-containing molecule is related to human complement factor H.		
#cross-references	MUID:93155112		

FEATURE	1-48	49-597	50-108	113-170	175-234	239-294	299-360	364-422	381-404	426-480	446-538	521,506,528	SUMMARY	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	524	525	526	527	528	529	530	531	532	533	534	535	536	537	538	539	540	541	542	543	544	545	546	547	548	549	550	551	552	553	554	555	556	557	558	559	560	561	562	563	564	565	566	567	568	569	570	571	572	573	574	575	576	577	578	579	580	581	582	583	584	585	586	587	588	589	590	591	592	593	594	595	596	597	598	599	600	601	602	603	604	605	606	607	608	609	610	611	612	613	614	615	616	617	618	619	620	621	622	623	624	625	626	627	628	629	630	631	632	633	634	635	636	637	638	639	640	641	642	643	644	645	646	647	648	649	650	651	652	653	654	655	656	657	658	659	660	661	662	663	664	665	666	667	668	669	670	671	672	673	674	675	676	677	678	679	680	681	682	683	684	685	686	687	688	689	690	691	692	693	694	695	696	697	698	699	700	701	702	703	704	705	706	707	708	709	710	711	712
---------	------	--------	--------	---------	---------	---------	---------	---------	---------	---------	---------	-------------	---------	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

Db	219	1HCADSKMNSPPACENPCINCPDIPAHSMWEIYPRPTKEDVYVGVLYARKYCHGKXYP	338
Qy	231	AVC-TGSGWNP-QSGCEMTCCLT-PLYNPGLY-T-PIRIRHRI--DDEI-RYCKNGKGFY	282
Db	339	TTDEPTVYICKNLRMPYOGCEMLACPEKLNNGEIT-OH-EKS-RPANHC-VY-FYGD	393
Qy	283	ATRSBPV-KCIT-T-GWIPAPRCLAKPDFQFKHGRLYEESNRPFYPPVIGKEIYSYCD	340
Db	394	EISFCHETSFFSAI-CQGDGTSMPRPSCGDICNP-PIAHGHYKQSSSYFFFEKEIT	451
Qy	341	N-GFTTSQSYWDYLRCOTVNG-WEPEVP-CLROCIFFHYVEGSSYSWQR-YIEGOSAKY	396
Db	452	YECDKGYLLV-GOAKLSCSYSHMVAAPDC	480
Qy	397	-QCHSGISLPNGQOTTYCTENGWSP-PRC	424
RESULT 12			
ENTRY	S57953 #type complete		
TITLE	C4B protein alpha chain precursor - rat		
ORGANISM	#formal_name Rattus norvegicus #common_name Norway rat		
DATE	13-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 20-Aug-1999		
ACCESSIONS	S57953		
REFERENCE	S57953		
#authors	Hillarp, A.; Thern, A.; Dahlback, B.		
#submission	Submitted to the EMBL Data Library, July 1995		
#description	Molecular cloning of rat C4b-binding protein alpha- and beta-chains: structural and functional relationships among human, bovine, rabbit, mouse and rat proteins.		
#accession	S57953		
#status	preliminary		
#molecule_type	mRNA		
#residues	1-558 #label HIL		
#cross-references	EMBL:Z50051; NID:g899379; PIDN:CAA90391.1; PID:g899380		
CLASSIFICATION	#superfamily C4b-binding protein alpha chain; complement factor H repeat homology		
FEATURE			
15-72	#domain complement factor H repeat homology #label FH1		
77-134	#domain complement factor H repeat homology #label FH2		
139-199	#domain complement factor H repeat homology #label FH3		
204-258	#domain complement factor H repeat homology #label FH4		
263-324	#domain complement factor H repeat homology #label FH5		
328-386	#domain complement factor H repeat homology #label FH6		
390-443	#domain complement factor H repeat homology #label FH7		
447-501	#domain complement factor H repeat homology #label FH8		
SUMMARY	#length 558 #molecular_weight 62465 #checksum 9928		
Query Match	14.7%: Score 494; DB 2; Length 558;		
Best Local Similarity	29.5%; Pred. No. 8.78e-89;		
Matches 132; Conservatave	89; Mismatches 182; Indels 45; Gaps 38;		
Db	15	C-GPPP-DLPALPASENNQTDIFESHITLRLNCRPGYSRRASSQSGLYCKPLGKW-QINIA	71
Qy	3	CKGPPRENSSEILSGSWSEQL-YSGGATATYKCRPGV-RTLGIIVYCKN-GEWVPNS	59
Db	72	--CVKSKCRNPDGLONG--KYEVKTDPLFGSQIEFSCSEGIILIGSSTSCIEIOGKGVSW	127
Qy	60	RICRRRPPGCHGDDTPFGSFRLAVSGSEFFGAKVYITDDEGYQLGGE-IDVRECDAGS--W	116
Db	128	SDPLRECIYACGMW-PIISNGK-HNGE-E-EEFFYRSSTVYKCDPDFLLGNASTICT	183
Qy	117	TNDIFICEVYKC-LPVTLENGRIYGAALPDQERYQVYRFECNSGFKIEGKEHMS	175
Db	184	VYVKTGVWSPBPPTCEBIIICPMWRVLAHTINSFKAHTYKXDSVRFVYCKGFYLRSGV	243
Qy	176	E-N--GLMSNKKPCQCVISICLIPRENGD-GIYLKPYKENEERFOYKCKQGFYKRGD	230
Db	244	IHCADSGMSPPVVELNSCIDPIDIPAAALITSPRPKKEVYVGVVLYARKYCHGKXYP	303
Qy	231	AVCTSG-WNDPSCSEETCL-TPYIPNG--IYHP-RIKH-RIDELIRECKNGGYPA	283



(WIL)

Distribution rights by Oxford Molecular Ltd

Tabular output not generated

Sequence:

Scoring table

Searched:

Statistics:

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	2834	84.1	1234	1	CEFA_MOUSE	0.00e+00
2	2396	71.1	1231	1	CEFA_HUMAN	0.00e+00
3	646	19.2	331	1	CEHD_HUMAN	8.26e-14
4	537	15.9	597	1	C4B_HUMAN	1.67e-11
5	494	14.7	558	1	C4B_RAT	1.05e-09
6	474	14.1	610	1	C4B_BOVIN	3.15e-04
7	446	13.2	2039	1	CR1_HUMAN	1.27e-86
8	434	12.9	330	1	CFH1_HUMAN	2.23e-83
9	434	12.9	469	1	C4B_MOUSE	2.23e-83
10	431	12.8	270	1	CFHE_HUMAN	1.44e-82
11	396	11.7	263	1	VCP_VACV	3.54e-73
12	390	11.6	830	1	LEM3_HUMAN	1.41e-71
13	375	11.1	345	1	APOR_MOUSE	1.37e-67
14	374	11.1	768	1	LEM3_RAT	2.51e-66
15	367	10.9	345	1	APOR_HUMAN	1.79e-66
16	368	10.9	768	1	LEM3_MOUSE	9.73e-66
17	364	10.8	345	1	APOR_CANFA	1.11e-66
18	357	10.6	610	1	LEM2_HUMAN	7.76e-65
19	359	10.6	612	1	LEM2_MOUSE	2.31e-65
20	348	10.3	377	1	F13B_MOUSE	1.80e-66
21	345	10.2	377	1	MCP_HUMAN	1.10e-55
22	344	10.2	1025	1	CR2_MOUSE	2.10e-55
23	341	10.1	345	1	APOR_BOVIN	1.23e-55
					BETA-2-GLYCOPROTEIN I	

24	335	9.9	769	1	LEM3_SHEEP	P-SELECTIN PRECURSOR	(4.52e-57
25	329	9.8	551	1	LEM2_RABBIT	E-SELECTIN PRECURSOR	(1.65e-55
26	328	9.7	507	1	DAF_CAVPO	COMPLEMENT DECAY-ACCEL	(3.00e-35
27	327	9.7	661	1	F13B_HUMAN	COAGULATION FACTOR XII	(3.47e-55
28	328	9.7	958	1	HIG_DROME	LOCOMOTION-RELATED PRO	(3.00e-55
29	323	9.6	384	1	LEM2_PIG	E-SELECTIN PRECURSOR	(5.97e-54
30	313	9.6	481	1	DAF_HUMAN	COMPLEMENT DECAY-ACCEL	(2.30e-51
31	309	9.2	297	1	APHO_RAT	BETA-2-GLYCOPROTEIN I	(2.47e-50
32	306	9.1	1033	1	CR2_HUMAN	COMPLEMENT RECEPTOR TY	(1.46e-49
33	304	9.0	611	1	LEM2_CANAFA	E-SELECTIN PRECURSOR	(4.77e-49
34	301	8.9	485	1	LEM2_BOVIN	E-SELECTIN PRECURSOR	(2.81e-48
35	294	8.7	340	1	DAF_PONY	COMPLEMENT DECAY-ACCEL	(1.73e-46
36	292	8.7	549	1	LEM2_RAT	E-SELECTIN PRECURSOR	(5.61e-46
37	281	8.6	390	1	DAF1_MOUSE	COMPLEMENT DECAY-ACCEL	(1.01e-45
38	276	8.2	407	1	DAF2_MOUSE	COMPLEMENT DECAY-ACCEL	(6.42e-42
39	274	8.1	360	1	CCPH_HSVIN	COMPLEMENT CONTROL PRO	(2.05e-41
40	268	8.0	646	1	LEM3_BOVIN	P-SELECTIN PRECURSOR	(6.61e-40
41	255	7.6	1019	1	LFC_TACR1	LIMDUS CLOTTING FACTO	(1.14e-36
42	252	6.9	372	1	LEM1_MOUSE	E-SELECTIN PRECURSOR	(5.26e-31
43	226	6.7	372	1	LEM1_PAPHA	L-SELECTIN PRECURSOR	(1.49e-29
44	226	6.7	372	1	LEM1_MACMU	L-SELECTIN PRECURSOR	(1.49e-29
45	225	6.7	372	1	LEM1_RALM	L-SELECTIN PRECURSOR	(2.60e-29

ALIGNMENTS

RESULT	1	STANDARD:	PRT: 1234 AA.
ID	CPAH MOUSE		
AC	P06909:		
DT	01-JAN-1988 (Rel. 06, Created)		
DT	01-JAN-1988 (Rel. 06, Last sequence update)		
DT	01-NOV-1997 (Rel. 35, Last annotation update)		
DE	COMPLEMENT FACTOR H PRECURSOR (PROTEIN BETA-1-H).		
GN	CFH.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;		
OC	Euteria; Rodentia; Sciurognathi; Muridae; Murine; Mus.		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE: 86233353.		
RA	Kristensen T., Tack B.F.;		
RT	"Murine protein H is comprised of 20 repeating units, 61 amino acids		
RT	in length".		
RL	Proc. Natl. Acad. Sci. U.S.A. 83:3963-3967(1986).		
RN	[2]		
RP	SEQUENCE OF 1-19 FROM N.A.		
RP	STRAIN-BALB/C.		
RC	MEDLINE: 90148935.		
RX	Munoz-Canoes P., Tack B.F., Vik D.P.;		
RA	"Analysis of complement factor H mRNA expression: dexamethasone and		
RT	IFN-gamma increase the level of H in L cells.";		
RT	Biochemistry 28:9891-9897(1989).		
RL	[3]		
RN	SEQUENCE OF 1-18 FROM N.A.		
RP	MEDLINE: 90111033.		
RX	Natlunne-Sekai S., Nonaka M., Harada Y.N., Shreffler D.C.,		
RA	Moriwaki K.;		
RT	"Demonstration of an unusual allelic variation of mouse factor H by		
RT	the complete cDNA sequence of the H.2 allele.";		
RL	J. Immunol. 144:358-362(1990).		
CC	-1- FUNCTION: FACTOR H FUNCTIONS AS A COFACTOR IN THE INACTIVATION OF		
CC	C3B BY FACTOR I AND ALSO INCREASES THE RATE OF DISSOCIATION OF THE		
CC	C3BB COMPLEX (C3 CONVERTASE) AND THE (C3)NBB COMPLEX (C5		
CC	CONVERTASE) IN THE ALTERNATIVE COMPLEMENT PATHWAY.		
CC	-1- POLYMORPHISM: TWO CODOMINANT ALLELES OF FACTOR H ARE PRESENT IN		
CC	MICE.		
CC	-1- SIMILARITY: CONTAINS 20 SUSHI (SCR) REPEATS.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on ways		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; M12660; AAA37759.1; -

DR EMBL; J02891; AAA37795.1; -

DR EMBL; M31979; AAA37762.1; -

DR PIR; A26154; NEMSH.

DR HSSP; P08603; IHFT.

DR MGD; MGI:88385; CFH.

DR PFAM; PF00084; sushi; 20.

KW Complement alternate pathway: Plasma; Glycoprotein; Repeat; Sushi; Signal.

FT SIGNAL 1 18

FT CHAIN 19 1234

FT DOMAIN 20 1224

FT REPEAT 20 81

FT REPEAT 84 142

FT REPEAT 145 206

FT REPEAT 209 263

FT REPEAT 266 321

FT REPEAT 324 386

FT REPEAT 388 443

FT REPEAT 447 506

FT REPEAT 508 565

FT REPEAT 568 623

FT REPEAT 628 684

FT REPEAT 689 744

FT REPEAT 751 803

FT REPEAT 807 862

FT REPEAT 866 935

FT REPEAT 935 990

FT REPEAT 993 1049

FT REPEAT 1052 1108

FT REPEAT 1113 1169

FT REPEAT 1171 1234

FT DISULFID 21 66

FT DISULFID 52 80

FT DISULFID 85 129

FT DISULFID 114 141

FT DISULFID 146 192

FT DISULFID 178 205

FT DISULFID 210 251

FT DISULFID 237 262

FT DISULFID 267 309

FT DISULFID 294 320

FT DISULFID 325 374

FT DISULFID 357 385

FT DISULFID 389 431

FT DISULFID 416 442

FT DISULFID 448 494

FT DISULFID 477 505

FT DISULFID 509 553

FT DISULFID 536 564

FT DISULFID 569 610

FT DISULFID 597 622

FT DISULFID 629 672

FT DISULFID 658 683

FT DISULFID 690 732

FT DISULFID 718 743

FT DISULFID 752 791

FT DISULFID 780 802

FT DISULFID 808 850

FT DISULFID 836 861

FT DISULFID 867 920

FT DISULFID 906 931

FT DISULFID 936 978

FT DISULFID 964 989

FT DISULFID 994 1037

FT DISULFID 1023 1048

FT DISULFID 1053 1096

FT DISULFID 1082 1107

FT DISULFID 1114 1157

FT DISULFID 1143 1168

BY SIMILARITY

COMPLEMENT FACTOR H.

20 X SUSHI (SCR) REPEATS.

SUSHI 1.

SUSHI 2.

SUSHI 3.

SUSHI 4.

SUSHI 5.

SUSHI 6.

SUSHI 7.

SUSHI 8.

SUSHI 9.

SUSHI 10.

SUSHI 11.

SUSHI 12.

SUSHI 13.

SUSHI 14.

SUSHI 15.

SUSHI 16.

SUSHI 17.

SUSHI 18.

SUSHI 19.

SUSHI 20.

FT DISULFID 1172 1223

FT DISULFID 1206 1233

FT CARBOHYD 676 676

FT CARBOHYD 721 721

FT CARBOHYD 773 773

FT CARBOHYD 801 801

FT CARBOHYD 1030 1030

FT CARBOHYD 1061 1061

FT CARBOHYD 1225 1225

FT CARBOHYD 1234 AA; 139082 MW; C5AC02F341B957F7 CRC64;

SO SEQUENCE

Query Match 84.1%; Score 2834; DB 1; Length 1234;

Best Local Similarity 83.0%; Pred. No. 0.00e+00;

Matches 356; Conservative 34; Mismatches 37; Indels 2; Gaps 2;

Db 19 EDCGPPPRENSEILSGSWSEQLYPEGTOATYKCRPGYRTLGIIVKCKNGKWNPSR 78

QY 1 EDCGPPPRENSEILSGSWSEQLYSEGTOATYKCRPGYRTLGIIVKCKNGKWNPSR 60

Db 79 ICRKPCGHPGDPFGSFRILAVGSOPEFGAKVYITCDGQVLLGEIDYRECGADGWTNDI 138

QY 61 ICRKPCGHPGDPFGSFRILAVGSEFEFGAKVYITCDGQVLLGEIDYRECGADGWTNDI 120

Db 139 PCEVVKCLPTELENGRIYSGAETDOEYFGQVYRFECSNGFKIEGHEICSENGI 198

QY 121 PCEVVKCLPTELENGRIYSGAETDOEYFGQVYRFECSNGFKIEGHEICSENGI 180

Db 199 SNKPRVVELCTPPRVENDGINVKRYKENDRYHKCKHGYVPERGDAVCTGSGWS 258

QY 181 SNKPRVVELCTPPRVENDGINVKRYKENDRYHKCKHGYVPERGDAVCTGSGWS 240

Db 259 QPCEKRCSPYILNGIYTPHRIHRSDEIRYECNMGYPPTGSKCTPTGWIIP 318

QY 241 QPCEKRCSPYILNGIYTPHRIHRSDEIRYECNMGYPPTGSKCTPTGWIIP 300

Db 319 RCLTKCEPFOFYGYLYEESLRPNFVPSIGNKYSKCDNGFSPPS-GYSMDYLCTAQ 377

QY 301 RCLTKCEPFOFYGYLYEESLRPNFVPSIGNKYSKCDNGFSPPS-GYSMDYLCTAQ 359

Db 378 GMEPEVPCRCQCFHVEYGESEYKORRYEGOSAKVVOCHSGYSLPNDOTYTCENGWS 437

QY 360 GMEPEVPCRCQCFHVEYGESEYKORRYEGOSAKVVOCHSGYSLPNDOTYTCENGWS 419

Db 438 PPKCIRIK 446

QY 420 PPKCIRIK 428

RESULT 2

ID CFAH_HUMAN STANDARD; PRT; 1231 AA.

AC P08603;

DT 01-AUG-1988 (Rel. 08, Created)

DI 01-JAN-1990 (Rel. 13, Last sequence update)

DE 15-FEB-2000 (Rel. 39, Last annotation update)

DE COMPLEMENT FACTOR H PRECURSOR.

GN HFI OR HF OR CFH.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

OC Eukaryota; Primates; Catarrhini; Hominoidea; Homo.

OC [1]

OC SEQUENCE FROM N.A.

RP TISSUE-LIVER.

RC MEDLINE; 88134059.

RA Ripocne J., Day A.J., Harris T.J.R., Sim R.B.;

RT "The complete amino acid sequence of human complement factor H.;"

RL Biochem. J. 249:593-602(1988).

RN [2]

RP SEQUENCE OF 53-445 FROM N.A.

RP MEDLINE; 87054207.

RA Schulz T.F., Schwaebel W., Stanley K.K., Weiss E., Dierich M.P.;

RT "Human complement factor H: Isolation of cDNA clones and partial cDNA

sequence of the 38-kDa tryptic fragment containing the binding site

for C3b.";

RL Eur. J. Immunol. 16:1351-1355(1986).
 RN [3]
 RP SEQUENCE OF 226-445 FROM N.A. AND PARTIAL SEQUENCE.
 RX MEDLINE: 86169701.
 RA Kristensen T., Wetzel R.A., Tack B.F.;
 RT "Structural analysis of human complement protein H: homology with C4b
 binding protein, beta 2-glycoprotein I, and the Ba fragment of B2.";
 RL J. Immunol. 136:3407-3411(1986).
 RN [4]
 RP SEQUENCE OF 1047-1231 FROM N.A.
 RX MEDLINE: 91201892.
 RA Esteller C., Koistinen V., Schwaible W., Dierich M.P., Weiss E.H.;
 RT "Cloning of the 1.4-kb mRNA species of human complement factor H
 reveals a novel member of the short consensus repeat family related
 to the carboxy terminal of the classical 150-kDa molecule.";
 RL J. Immunol. 146:3190-3196(1991).
 RN [5]
 RP SEQUENCE OF 19-35.
 RX MEDLINE: 83048213.
 RA Sim R.B., Discipio R.G.;
 RT Purification and structural studies on the complement-system control
 protein beta 1H (factor H).";
 RL Biochem. J. 205:285-293(1982).
 RN [6]
 RP STRUCTURE BY NMR OF 927-985 (SUSHI 16).
 RX MEDLINE: 91278097.
 RA Norman D.G., Barlow P.N., Baron M., Day A.J., Sim B., Campbell I.D.;
 RT "Three-dimensional structure of a complement control protein module
 in solution.";
 RL J. Mol. Biol. 219:717-725(1991).
 RN [7]
 RP STRUCTURE BY NMR OF 264-322 (SUSHI 5).
 RX MEDLINE: 92232649.
 RA Barlow P.N., Norman D.G., Steinkasserer A., Horne T.J., Pearce J.,
 Driscoll P.C., Sim B., Campbell I.D.;
 RT "Solution structure of the fifth repeat of factor H: a second example
 of the complement control protein module.";
 RL Biochemistry 31:3626-3634(1992).
 RN [8]
 RP STRUCTURE BY NMR OF 866-985 (SUSHI 15 AND 16).
 RX MEDLINE: 93323119.
 RA Barlow P.N., Steinkasserer A., Norman D.G., Kieffer B., Wiles A.P.,
 Sim B., Campbell I.D.;
 RT "Solution structure of a pair of complement modules by nuclear
 magnetic resonance.";
 RL J. Mol. Biol. 232:268-284(1993).
 CC -1- FUNCTION: FACTOR H FUNCTIONS AS A COPROTECTOR IN THE INACTIVATION OF
 C3B BY FACTOR I AND ALSO INCREASES THE RATE OF DISSOCIATION OF THE
 C3BB COMPLEX (C3 CONVERTASE) AND THE (C3B)NBB COMPLEX (C3
 CONVERTASE) IN THE ALTERNATIVE COMPLEMENT PATHWAY.
 CC -1- SIMILARITY: CONTAINS 20 SUSHI (SCR) REPEATS.
 CC -1- CAUTION: REP. 2 SEQUENCE DIFFERS FROM THAT SHOWN FROM POSITION
 341 ONWARD DUE TO A FRAMESHIFT.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: Y00716; CA68704.1; -
 DR EMBL: X04697; CAB41739.1; ALT_FRAME.
 DR EMBL: M65294; AAA35948.1; -
 DR PIR: S00254; NBRUH.
 DR PIR: S00254; S00254.
 DR PIR: S03013; S03013.
 DR PDB: 1HCC; 15-APR-92.
 DR PDB: 1HEH; 15-JUL-93.
 DR PDB: 1HEH; 15-JUL-93.
 DR MIM: 134370; -
 DR PFAM: PF00084; sushi; 20.

KW Complement alternate pathway; Plasma; Glycoprotein; Repeat; Sushi;
 KM Signal; 3D-structure; Polymorphism.
 FT CHAIN 1 18
 FT SIGNAL 1 131
 FT DOMAIN 20 1229
 FT REPEAT 20 81
 FT REPEAT 84 142
 FT REPEAT 145 206
 FT REPEAT 209 263
 FT REPEAT 266 321
 FT REPEAT 324 386
 FT REPEAT 388 443
 FT REPEAT 447 506
 FT REPEAT 508 565
 FT REPEAT 568 624
 FT REPEAT 629 685
 FT REPEAT 690 745
 FT REPEAT 752 804
 FT REPEAT 810 865
 FT REPEAT 869 927
 FT REPEAT 930 985
 FT REPEAT 988 1044
 FT REPEAT 1047 1103
 FT REPEAT 1108 1164
 FT REPEAT 1166 1229
 FT DISULFID 21 66
 FT DISULFID 52 80
 FT DISULFID 85 129
 FT DISULFID 114 141
 FT DISULFID 146 192
 FT DISULFID 178 203
 FT DISULFID 210 251
 FT DISULFID 237 262
 FT DISULFID 267 309
 FT DISULFID 294 320
 FT DISULFID 325 374
 FT DISULFID 357 385
 FT DISULFID 389 431
 FT DISULFID 416 442
 FT DISULFID 448 494
 FT DISULFID 477 505
 FT DISULFID 509 553
 FT DISULFID 536 564
 FT DISULFID 569 611
 FT DISULFID 597 623
 FT DISULFID 630 673
 FT DISULFID 659 684
 FT DISULFID 691 744
 FT DISULFID 719 733
 FT DISULFID 753 792
 FT DISULFID 781 803
 FT DISULFID 811 853
 FT DISULFID 839 864
 FT DISULFID 870 915
 FT DISULFID 901 926
 FT DISULFID 931 973
 FT DISULFID 959 984
 FT DISULFID 989 1032
 FT DISULFID 1018 1043
 FT DISULFID 1048 1091
 FT DISULFID 1077 1102
 FT DISULFID 1109 1152
 FT DISULFID 1138 1163
 FT DISULFID 1167 1218
 FT DISULFID 1201 1228
 FT CARBOHYD 529 529
 FT CARBOHYD 718 718
 FT CARBOHYD 802 802
 FT CARBOHYD 822 822
 FT CARBOHYD 862 862
 FT CARBOHYD 911 911
 FT CARBOHYD 1029 1029
 FT CARBOHYD 1095 1095

COMPLEMENT FACTOR H.
 20 X SUSHI (SCR) REPEATS.

POTENTIAL.

POTENTIAL.

```

FT VARIANT 402 402 H -> Y.
FT CONFLICT 21 21 /FTID=VAR_001979.
FT CONFLICT 21 21 C -> Q (IN REF. 3).
FT CONFLICT 30 30 T -> V (IN REF. 3).
FT CONFLICT 34 34 T -> Q (IN REF. 3).
FT CONFLICT 53 54 RP -> IL (IN REF. 2).
FT STRAND 870 871
FT STRAND 876 876
FT TURN 877 878
FT STRAND 879 881
FT STRAND 890 891
FT TURN 893 894
FT STRAND 896 900
FT TURN 903 904
FT STRAND 907 907
FT STRAND 912 916
FT TURN 917 918
FT STRAND 919 920
FT STRAND 926 926
SQ SEQUENCE 1231 AA; 139125 MW; C65EC8CF8800B3FD CRC64;

```

```

Query Match 71.1%; Score 2396; DB 1; Length 1231;
Best Local Similarity 67.5%; Pred. No. 0.00e+00;
Matches 289; Conservative 64; Mismatches 75; Indels 0; Gaps 0;

```

```

Db 19 EDCNMLPRRRTTLLTSGMSDQTYPEGTQAIYKCRPGTRSLGNVTMCRKGEWALNPLR 78
| 1 EDCGPPPRENSEILSGMSDQYSEGTQAIYKCRPGTRTIGTIYKVCNGEWMVNSPSR 60
| 79 KCRKPCGHPDPTFGFTLTGNGVFEYGVAVYTCNGEYOLGELNYRECDTGWTNDI 138
| 61 ICRKRPCHPDDTPGSGRLAVGSEFEFGAVYTCDSQYOLGELIDRECDADGWTNDI 120
| 139 PICVYKCLPTAENKGVSSAMEPDRYHFGQAVRVCNSGKYEDEMGSDDEFTW 198
| 121 PICVYKCLPTAENKGVSSAMEPDRYHFGQAVRVCNSGKYEDEMGSDDEFTW 180
| 199 SKERPKVEICSPDYVNGSPISQIYKEMERPOYKCNNGYESEGDVAVCESGWRP 258
| 181 SNERPCVEICSPDYVNGSPISQIYKEMERPOYKCNNGYESEGDVAVCESGWRP 240
| 259 LPSCSEKSCNDPIYVNGSDYPLRIKHTGDEITYOCNGEFPATRGNTAKCTSGWIPAP 318
| 241 QPCEEMTCLTPYIPNGIYTPRIKHRIIDREYCKNGEFPATRSYKCTITIGWIPAP 300
| 319 RCTKPCDYPDIKGLHGMRRPVPVAVGKTYSTYCDHEFTPSGSYWDHICTQDG 378
| 301 RCTKPCDYPDIKGLHGMRRPVPVAVGKTYSTYCDHEFTPSGSYWDHICTQDG 360
| 379 WSPAVPCLRCYFYLENGYNQNGRKFGVQKSIDVACHPGLAIPKQTYTCMENGMS 438
| 361 WEPEVPCLRQICFHYVYEGSSYQWRRIEYQSAKYCHSGISLIPNGODITYCTENGMS 420
| 439 TPRCIRK 446
| 421 PRCVIRK 428

```

```

RESULT 3
ID CEHD_HUMAN STANDARD; PRT; 331 AA.
AC 002985;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DE 01-JUN-1994 (Rel. 29, Last annotation update)
DE COMPLEMENT FACTOR H-LIKE PROTEIN DOWN6 PRECURSOR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Homiidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE: 9315512.
RA Skerka C., Kuehn S., Guenther K., Lingelbach K., Zipfel P.F.;

```

```

FT "A novel short consensus repeat-containing molecule is related to
RT human complement factor H."
J. Biol. Chem. 268:2904-2908(1993).
CC -1- FUNCTION: MIGHT BE INVOLVED IN COMPLEMENT REGULATION.
CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC -1- SIMILARITY: CONTAINS 5 SUSHI (SCR) REPEATS.
CC -1- SIMILARITY: STRONG, TO FACTOR H.

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).

```

DR EMBL: X68679; CAA48639.1; -

DR PIR: A45222; A45222.

DR HSSP: P10998; 1VVC.

DR PFAM: PF00084; sush1; 5.

KW Repeat; Glycoprotein; Sush1; Signal.

FT SIGNAL 1 18

FT CHAIN 19 331

FT DOMAIN 19 331

FT REPEAT 19 85

FT REPEAT 86 128

FT REPEAT 129 207

FT REPEAT 208 265

FT REPEAT 266 331

FT CARBOHYD 108 108

FT CARBOHYD 186 186

FT CARBOHYD 206 206

FT CARBOHYD 310 310

FT SEQUENCE 331 AA; 37469 MW; 3DC48293FD91332C CRC64;

Query Match 19.2%; Score 646; DB 1; Length 331;

Best Local Similarity 56.3%; Pred. No. 8.26e-142;

Matches 76; Conservative 26; Mismatches 32; Indels 1; Gaps 1;

Db 9 LTLWVSCANGVRCDDPDRIKGLHGMRRPVPVAVGKTYSTYCDHEFTPSGSYWD 68

QY 293 ITGWIPAPRCSLKPDPQKRGHGLYEESRRPFPVDPICKEYSYCDNGFTTSPQS YWD 352

Db 69 YHCTQWSPVAVPCLRCYFYLENGYNQNGRKFGVQKSIDVACHPGLAIPKQTY 128

QY 353 YLRCTVNGMEPEVPCLRQICFHYVYEGSSYQWRRIEYQSAKYCHSGISLIPNG-ODTY 411

Db 129 TCTENGSPTPRCIR 143

QY 412 YCTENGSPTPRCIR 426

```

RESULT 4
ID C4BP_HUMAN STANDARD; PRT; 597 AA.
AC P04003;
DT 23-OCT-1986 (Rel. 02, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE C4B-BINDING PROTEIN ALPHA CHAIN PRECURSOR (PROLINE-RICH PROTEIN)
DE (PRP).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Homiidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE: 90073699.
RA Matsuguchi T., Okamura S., Aso T., Sata T., Niho Y.;
RT "Molecular cloning of the cDNA coding for proline-rich protein (PRP);
RT identity of PRP as C4b-binding protein."
RL Biochem. Biophys. Res. Commun. 165:138-144(1989).
RN [2]

```


Db 219 VENETIGVWBPSPPTCEKTRKPDVSHGEMVSGFPIYKNTIVRCOKGFVLRGSSV 278
 QY 176 -EN---GLMSNEKQCVETISCLPRVENGDDY-LKPYKNEFFQYKCKOGFYKKEKRGD 230
 Db 279 IHCDDADSKNMPSPACPNPCINLPDIPHASWETYPRTEDVYVGVTLRYRCHPGYKP 338
 QY 231 AVC-TGSGWNP-QPSCHEMTCLT-PYIPNGIY-T-PRIKHRI--DDEL-RYECKNGFYF 282
 Db 339 TIDEPTVICOKNLRMTPYOGCEALCCPEKNGEIT-QH-RKS-RPANHC-VY-FYCD 393
 QY 283 ATSPYS-KCITL-GWIPARCSLPCDFQFQKGRLYEESRRPYVPVIGKEYSYCD 340
 Db 394 EISFSCHESTRFSAI-CQGDGTWSPRTSPGCDICNFP-PIAHGHYKSSSYSPFKEEII 451
 QY 341 N-GFTTSPQSYMDYLACTVNG-WEPEVP-CLROCIHFHYVEGSSYWR-R-YIEGOSAKY 396
 Db 452 YEGDKGYILV-GQAKLSCSYSHWSAPAPOC 480
 QY 397 -OCHSGTSLPNGODTYCTENGWSP-PRC 424

RESULT 5 STANDARD: PRT; 558 AA.

AC C4BP-RAT 063514:
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE C4B-BINDING PROTEIN ALPHA CHAIN PRECURSOR.
 GN C4BPA.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SPRAGUE-DAWLEY; TISSUE=LIVER;
 RX MEDLINE; 97166082.
 RA Hillarp A., Wiklund H., Thern A., Dahlback B.,
 RT "Molecular cloning of rat C4b binding protein alpha- and beta-chains:
 RT structural and functional relationships among human, bovine, rabbit,
 RT mouse, and rat proteins.";
 RL J. Immunol. 158:1315-1323(1997).
 CC -1- FUNCTION: C4BP CONTROLS THE CLASSICAL PATHWAY OF COMPLEMENT
 CC ACTIVATION. IT BINDS AS A COFACTOR TO C3B/C4B INACTIVATOR
 CC (C3BINA), WHICH THEN HYDROLYSES THE COMPLEMENT FRAGMENT C4B. IT
 CC ALSO ACCELERATES THE DEGRADATION OF THE C4B C2A COMPLEX (C3
 CC CONVERTASE) BY DISSOCIATING THE COMPLEMENT FRAGMENT C2A. ALPHA
 CC CHAIN BINDS C4B. IT INTERACTS ALSO WITH ANTICOAGULANT PROTEIN S
 CC AND WITH SERUM AMYLOID P COMPONENT.
 CC -1- SUBUNIT: DISULFIDE-LINKED COMPLEX OF C4BP ALPHA AND BETA CHAINS.
 CC -1- SIMILARITY: TO C4BP BETA CHAIN AND TO PIG APOLIPOPROTEIN R.
 CC -1- SIMILARITY: CONTAINS 8 SUSHI (SCR) REPEATS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by, and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Z50051; CAA90391.1; -.
 DR HSSP; P10998; IYVC.
 DR PFM; PFO0084; SUSHI; 8.
 KW Complement pathway; Plasma; Glycoprotein; Repeat; Sushi; Signal.
 FT CHAIN 1 13
 FT SIGNAL 13
 FT DOMAIN 14 558
 FT 14 502
 FT REPEAT 14 73
 FT REPEAT 14 76
 FT REPEAT 138 200
 FT REPEAT 203 259
 FT REPEAT 262 325
 FT SUSHI 5.

FT REPEAT 328 387
 FT REPEAT 389 444
 FT REPEAT 446 502
 FT DISULFID 15 60
 FT DISULFID 45 72
 FT DISULFID 77 118
 FT DISULFID 104 134
 FT DISULFID 139 182
 FT DISULFID 168 199
 FT DISULFID 204 246
 FT DISULFID 232 288
 FT DISULFID 263 312
 FT DISULFID 296 324
 FT DISULFID 328 351
 FT DISULFID 332 363
 FT DISULFID 363 431
 FT DISULFID 417 443
 FT DISULFID 447 488
 FT DISULFID 474 501
 FT DISULFID 509 509
 FT DISULFID 521 521
 FT CARBOHYD 31 31
 FT CARBOHYD 177 177
 FT CARBOHYD 186 186
 FT CARBOHYD 469 469
 FT CARBOHYD 491 491
 SQ SEQUENCE 558 AA; 62266 MW; 592F0C67ED1E5FF CRC64;

Query Match 14.7%; Score 494; DB 1; Length 558;
 Best Local Similarity 29.5%; Pred. No. 1,096-99;
 Matches 132; Conservative 89; Mismatches 182; Indels 45; Gaps 38;

Db 15 C-GPP-DLPYALPASEMANGDDESHHTLRNCRPGYRASSSSOSLYCKPIKQ-OINIA 71
 QY 3 CKGPPRENSSETLSSGWSSEOL-YSEGTQATYKCRPGY-RTIGTIYKVKCN-GEWVSPNS 59
 Db 72 -CVKSCRNFGDLDNG--KYEVKTDLPFGSOIEPSCSEGYILGSSSTSYCEIGKGYSW 127
 QY 60 RICRRAPCGHGDTPFGSFRILAVGSEFFGAKVYITCEGYQLGE-IDYECDDAG--W 116
 Db 128 SDLPPECYIACMP-PDISNGK-HNGR-E-EFFTYRASYTYKCDPFTLLGNASITCT 183
 QY 117 TNDIPICVYVC-LPVELLENGRIYSGAEPDOEYFGQVAVFPCNSGFKIEGKEMHS 175
 Db 184 VYNKVGYWSPSPCEHICLPMKVLHGITNSGFKHYKDKSVRYVCOKGFVLRGSGV 243
 QY 176 E-N---GLMSNEKQCVETISCLPRVENGD-GIYLKPYKNEFFQYKCKOGFYKKEKRGD 230
 Db 244 IHCDAWSMSPVPCVCEINSCDIPDIPNALITSPRKEDEVYVGVTLRYICRGPYEA 303
 QY 231 AVCTGSG-WNPQSCHEMTCL-TYIPNG--ITYPH-RIKH--RIDDIRKCNNGFYPA 283
 Db 304 TROMTYICQDLWSMLRGKEICCVPPDKSVYI-QHEK-AH-P-DNDCTY-FRGDE 358
 QY 284 TRSVVS-KCITL-GWIPARCSLPCDFQFQKGRLYEESRRPYVPVIGKEYSYCDN 341
 Db 359 VSYTCNDIMLT-ATCKSDGWHMPTSPCHOSODFPAIHAHGRTRKSSYVRYQV-YE 416
 QY 342 -GFTTSPQSYMDYLACTVNG-WEPEVP-CLROCIHF-YVEGSSYWRRIEGOSAKY 397
 Db 417 CEESYRLV-GEATISCMYSQWTPAAPOC 443
 QY 398 CHSGTSLPNGODTYCTENGWSP-PRC 424

RESULT 6 STANDARD: PRT; 610 AA.
 AC Q28065;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE C4B-BINDING PROTEIN ALPHA CHAIN PRECURSOR.

GN CABPA.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae;
 OC Bovinae; Bos.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-LIVER.
 RX MEDLINE: 95015909.
 RA Hillarp A., Thern A., Dahlbaeck B.;
 RT "Bovine C4b binding protein. Molecular cloning of the alpha- and
 beta-chains provides structural background for lack of complex
 formation with protein S.";
 RL J. Immunol. 153:4190-4199(1994).
 CC -1- FUNCTION: C4BP CONTROLS THE CLASSICAL PATHWAY OF COMPLEMENT
 ACTIVATION. IT BINDS AS A COFACTOR TO C3b/C4b INACTIVATOR
 (C3bINA), WHICH THEN HYDROLYZES THE COMPLEMENT FRAGMENT C4b. IT
 ALSO ACCELERATES THE DEGRADATION OF THE C4bC2a COMPLEX (C3
 CONVERTASE) BY DISSOCIATING THE COMPLEMENT FRAGMENT C2a. ALPHA
 CHAIN BINDS C4b. IT INTERACTS ALSO WITH SERUM AMYLOID P COMPONENT.
 CC -1- SUBUNIT: DISULFIDE-LINKED COMPLEX OF C4BP ALPHA AND BETA CHAINS
 (BY SIMILARITY).
 CC -1- SIMILARITY: TO C4BP BETA CHAIN AND TO PIG APOLIPOPROTEIN R.
 CC -1- SIMILARITY: CONTAINS 8 SUSHI (SCR) REPEATS.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: Z31693; CAA83498.1; -
 DR HSSP: P10998; IYVC.
 DR PFAM: PF00084; SUSHI; 8.
 DR Complement pathway; Plasma; Glycoprotein; Repeat; Sushi; Signal.
 KW SIGNAL
 FT CHAIN 1 48 BY SIMILARITY.
 FT DOMAIN 49 610 C4b-BINDING PROTEIN ALPHA CHAIN.
 FT REPEAT 49 108 8 X SUSHI (SCR) REPEATS.
 FT REPEAT 111 170 SUSHI 1.
 FT REPEAT 173 235 SUSHI 2.
 FT REPEAT 238 295 SUSHI 3.
 FT REPEAT 298 363 SUSHI 4.
 FT REPEAT 366 426 SUSHI 5.
 FT REPEAT 428 484 SUSHI 6.
 FT REPEAT 486 542 SUSHI 7.
 FT REPEAT 542 595 SUSHI 8.
 FT DISULFID 50 95 BY SIMILARITY.
 FT DISULFID 80 107 BY SIMILARITY.
 FT DISULFID 112 153 BY SIMILARITY.
 FT DISULFID 139 169 BY SIMILARITY.
 FT DISULFID 174 217 BY SIMILARITY.
 FT DISULFID 203 234 BY SIMILARITY.
 FT DISULFID 239 281 BY SIMILARITY.
 FT DISULFID 267 294 BY SIMILARITY.
 FT DISULFID 299 350 BY SIMILARITY.
 FT DISULFID 334 362 BY SIMILARITY.
 FT DISULFID 336 390 BY SIMILARITY.
 FT DISULFID 367 412 BY SIMILARITY.
 FT DISULFID 402 425 BY SIMILARITY.
 FT DISULFID 429 471 BY SIMILARITY.
 FT DISULFID 457 483 BY SIMILARITY.
 FT DISULFID 487 528 BY SIMILARITY.
 FT DISULFID 514 541 BY SIMILARITY.
 FT DISULFID 549 561 INTERCHAIN (WITH BETA CHAIN) (POTENTIAL).
 FT DISULFID 561 561 INTERCHAIN (WITH BETA CHAIN) (POTENTIAL).
 FT CARBOHYD 66 66 POTENTIAL.
 FT CARBOHYD 221 221 POTENTIAL.
 FT CARBOHYD 525 525 POTENTIAL.
 FT CARBOHYD 602 602 POTENTIAL.
 SQ SEQUENCE 610 AA: 68886 MW: D806B270E8A06B58 CRC64;

Query Match 14.18, Score 474, DB 1, Length 610;
 Best Local Similarity 27.9%, Pred. No. 3,15e-94;
 Matches 120; Conservative 85; Mismatches 180; Indels 45; Gaps 35;
 Db 70 FENGTLLRYCRRPGYRIRSSRNFLICDGTDNW--KYNE-FCVKKRCNPGEL-LNG-OVI 124
 QY 24 YSEGTQATYKCRPGYRIRLUG-TIVKVCNKE-WVPSNPSRCRKRPGCHGDDTPPGSRFLA 81
 Db 125 VKTDSFGSEIEFSCSEGYVLIGSANSYCOLQDKGVWSDPLPQCIIAKCEPPPTISNGR 184
 QY 82 VGSFEFGAKYVYTCDSGYQLGLGID-YRECDADG--WTNDIPICEVYKCLPVELENGR 138
 Db 185 -HNGGDE-DF-YTYGSSVYTSQDPSMLGKASISCRVENKTIWVSPSPSCKKVYCVQ 241
 QY 139 IVSGAAEPDQYRYGVGVVREFCNSGFKIEGQKEMKNS-EN--GLWSNKPQCVIEISCLP 194
 Db 242 PVMDKRTISGFGPIIYTYQOSIYVACNKGPELGGDSLHCADNSMNPPTCEINGCLG 301
 QY 195 PRVNGDGIY-LKPYIKENEFQYKCKGQFYKRGDAVC-TGSGMNPQ--SCEMTCIT 251
 Db 302 LPHIPALMERYDHQOTQDQVYDIGFVLSYKCHFGYKPEPDGPTTYVQSNLEMSPLYE 361
 QY 252 -PIYDNGI---YTPH-RIKHRI-DDE-I-RYCKNGFYPATRSPVS-KC-TITGWIAPAR 301
 Db 362 CKVCCPEPNLNTNGSITLH--RRP-----STSTCTYISGKISTECHKYMPDAL-CTK 414
 QY 302 CSLKCDPQPKRH-GRLYEESRRPFPVPGKEVSYCDNGFTTPSOS-Y-WDYLKCTV 358
 Db 415 HGTWSPRTPECRPDCKSPVIAHGKIVSKFTFDHQAIVECDKGYLLGAKELTS-CTS 473
 QY 359 NG-MEPEVP-CLRQCFIH-YVEYGESSYWRRIYEGSAKQVCHSGYSLPNGODIYVCTE 415
 Db 474 SGWSPAVPQC 483
 QY 416 NGWSP-PRC 424
 RESULT 7
 ID CRI-HUMAN STANDARD; PRT: 2039 AA.
 AC P17927;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE COMPLEMENT RECEPTOR TYPE 1 PRECURSOR (C3b/C4b RECEPTOR) (CD35
 DE ANTIGEN).
 DE CRI OR C3bR.
 OS Homo sapiens (Human).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Homnidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 89035992.
 RA Klinkstein L.B., Bartow T.J., Miletic V., Rabson L.D., Smith J.A.,
 RA Fearon D.T.;
 RT "Identification of distinct C3b and C4b recognition sites in the
 human C3b/C4b receptor (CRI, CD35) by deletion mutagenesis.";
 RL J. Exp. Med. 168:1699-1717(1988).
 RN [2]
 RP SEQUENCE OF 503-2039 FROM N.A.
 RX MEDLINE: 87168191.
 RA Klinkstein L.B., Wong W.W., Smith J.A., Weis J.H., Wilson J.G.,
 RA Fearon D.T.;
 RT "Human C3b/C4b receptor (CRI). Demonstration of long homologous
 repeating domains that are composed of the short consensus repeats
 RT characteristics of C3/C4 binding proteins.";
 RL J. Exp. Med. 165:1095-1112(1987).
 RN [3]
 RP SEQUENCE OF 761-783; 831-845 AND 1179-1195 FROM N.A.
 RX MEDLINE: 86067975.
 RA Wong W.W., Klinkstein L.B., Smith J.A., Weis J.H., Fearon D.T.;
 RT "Identification of a partial cDNA clone for the human receptor for
 RT complement fragments C3b/C4b.";
 RL Proc. Natl. Acad. Sci. U.S.A. 82:7711-7715(1985).

RT of the serum complement system."
 RL Biochemistry 26:4668-4674(1987).
 CC -1- FUNCTION: C4BP CONTROLS THE CLASSICAL PATHWAY OF COMPLEMENT
 CC ACTIVATION. IT BINDS AS A COFACTOR TO C3B/C4B INACTIVATOR
 CC (C3BIWA), WHICH THEN HYDROLYSES THE COMPLEMENT FRAGMENT C4B. IT
 CC ALSO ACCELERATES THE DEGRADATION OF THE C4Bc2a COMPLEX (C3
 CC CONVERTASE) BY DISSOCIATING THE COMPLEMENT FRAGMENT C2a. ALPHA
 CC CHAIN BINDS C4B. IT INTERACTS ALSO WITH SERUM AMYLOID P COMPONENT.
 CC -1- SUBUNIT: HOMOEPTAMER; NOT COVALENTLY LINKED. MOUSE LACKS THE
 CC BETA CHAIN OF C4BP.
 CC -1- SIMILARITY: TO C4BP BETA CHAIN AND TO PIG APOLIPOPROTEIN R.
 CC -1- SIMILARITY: CONTAINS 6 SUSHI (SCR) REPEATS.
 CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-44 IS THE INITIATOR.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M17122; AAA37312.1; ALT_INIT.
 DR PIR: A27117; NEMSC4.
 DR HSSP: P10998; IYVC.
 DR MGD: MGI:88329; C4BP.
 DR PFW: PF00084; sush1; 6.
 KW Complement pathway; Plasma; Glycoprotein; Repeat; Sush1; Signal.
 FT SIGNAL 1 56
 FT CHAIN 57 469 CAB-BINDING PROTEIN.
 FT DOMAIN 57 414 6 X SUSHI (SCR) REPEATS.
 FT REPEAT 57 116 SUSHI 1.
 FT REPEAT 119 177 SUSHI 2.
 FT REPEAT 180 241 SUSHI 3.
 FT REPEAT 244 300 SUSHI 4.
 FT REPEAT 302 356 SUSHI 5.
 FT REPEAT 358 414 SUSHI 6.
 FT DISULFID 58 103 BY SIMILARITY.
 FT DISULFID 88 115 BY SIMILARITY.
 FT DISULFID 120 160 BY SIMILARITY.
 FT DISULFID 146 176 BY SIMILARITY.
 FT DISULFID 181 223 BY SIMILARITY.
 FT DISULFID 209 240 BY SIMILARITY.
 FT DISULFID 245 287 BY SIMILARITY.
 FT DISULFID 273 299 BY SIMILARITY.
 FT DISULFID 303 343 BY SIMILARITY.
 FT DISULFID 329 355 BY SIMILARITY.
 FT DISULFID 359 400 BY SIMILARITY.
 FT DISULFID 386 413 BY SIMILARITY.
 FT CARBOHYD 74 74 POTENTIAL.
 FT CARBOHYD 227 227 POTENTIAL.
 FT CARBOHYD 275 275 POTENTIAL.
 FT CARBOHYD 292 292 POTENTIAL.
 FT CARBOHYD 366 366 POTENTIAL.
 FT CARBOHYD 381 381 POTENTIAL.
 FT CARBOHYD 428 428 POTENTIAL.
 SQ SEQUENCE 469 AA; 51551 MW; 418137CB8D8C6321 CRC64;
 Query Match 12.9%; Score 434; DB 1; Length 469;
 Best Local Similarity 31.0%; Pred. No. 2,236-83;
 Matches 116; Conservative 83; Mismatches 139; Indels 41; Gaps 36;
 Db 58 C-GPPP-ALPNALPASDVNRITDESHTLTLYCLDPYGGISRMVYVYCPSEGEWLSV-S 114
 QY 3 CKGPPREMSILLSSMSQQL-YSEGTOATYKCRPGY-RTLTITVYKVC-NSEWPSNPS 59
 Db 115 -CAKHCNPNLYLNG-Y-VN-GETTFGSQLIEFSQEGFLVSSSTSSCEVRGKYAM 169
 QY 60 RICRRRPGCHPDTPFGSRLAVGSEFEFGARVYTCDEGYQLDEIDYR-ECADAG--W 116
 Db 170 SNPEACTYKCGPPDDISNGK-HSGT-E-DF-YEYNGISTCOPGRVLVSSPTIGCTV 225
 QY 117 TNDIPICVAVCLPTELENGRIYGAAPDDEYFGVAVRECSNGFKIEQKEMHCSE 176

Db 226 VNKTVPWSSPPECKICSGPNILHGVISYKATYTHRDSVRLACNGTVLRGHVI 285
 QY 177 -N-GL-WSNKKPCVCEISCLPVPVENGDGY-LKPYKENEROYKOGFYKKEGDA 231
 Db 286 ECGNGWSSLPCE-EDCDLPPALVNGYYS-WYISKIT-LTYIEDCKGRLVKAIIIS 342
 QY 232 VCTSGS-WNPSPCEBMTCLTPYIPNGIYPRIKRIDEIRYECKNFEYPATRSPTS 289
 Db 343 -CSFSKMGTAPOCKA-LCKPVEVNGTLLS-DE-KDY--VE-SENATYICDSGFALMG- 394
 QY 290 KCTITGMV-PAPRCSLKPCEPPOCKHRLYVESRRYFPVPICKEISYICDNGFTTSPQ 348
 Db 395 S-OSISCSSEGTWYPEVPRC 413
 QY 349 SYWYLLCTVNG-WEPEVP-C 367
 RESULT 10
 ID CFEH_HUMAN STANDARD; PRT; 270 AA.
 AC P36980;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 01-JUN-1994 (Rel. 29, Last annotation update)
 DE COMPLEMENT FACTOR H-LIKE PROTEIN DDESK59 PRECURSOR.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Homiidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER;
 RX MEDLINE; 92251200.
 RA Skerka C., Timman C., Horstmann R.D., Zipfel P.F.;
 RT Two additional human serum proteins structurally related to
 RT complement factor H. Evidence for a family of factor H-related
 RT genes.
 RL J. Immunol. 148:3313-3318(1992).
 CC -1- FUNCTION: MIGHT BE INVOLVED IN COMPLEMENT REGULATION.
 CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.
 CC -1- SIMILARITY: CONTAINS 5 SUSHI (SCR) REPEATS.
 CC -1- SIMILARITY: STRONG, TO FACTOR H.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X64877; CAA46096.1; -
 DR PIR: S24564; S24564.
 DR HSSP: P08603; IHP1.
 DR PFW: PF00084; sush1; 4.
 KW Repeat; Glycoprotein; Sush1; Signal.
 FT SIGNAL 1 18
 FT CHAIN 19 270 POTENTIAL.
 FT DOMAIN 19 270 4 X SUSHI (SCR) REPEATS.
 FT REPEAT 19 85 SUSHI 1.
 FT REPEAT 86 127 SUSHI 2.
 FT REPEAT 128 205 SUSHI 3.
 FT REPEAT 206 270 SUSHI 4.
 FT CARBOHYD 126 126 POTENTIAL.
 SQ SEQUENCE 270 AA; 30651 MW; E1B2E9F139B217A8 CRC64;
 Query Match 12.8%; Score 431; DB 1; Length 270;
 Best Local Similarity 42.4%; Pred. No. 1,446-82;
 Matches 50; Conservative 22; Mismatches 46; Indels 0; Gaps 0;
 Db 23 CDFPKINGIYDEKTYFQSVPTGEVYFYSCYENFYSPKSFMTRTTCAEBSMSPTPK 82
 QY 307 CDFPKFKHGRILYBESRRPFPVPIGKEYIYICDNGFTTSPQSTWIDYLRCTVNGWEPEVP 366

Db 277 HSAKAFHOSCSFSCSEGFALVGP-EVVOCTASGWTAPAPVCKAVOCCHLEAPSEG-T 334
 QY 81 AVGSFEEFGARVYVTCDEGYQLGEIDRECDADG-WNDIPICEVYKCLPVTLENGRI 139
 Db 335 MD-CVHPITLTAAYGSGSCFECOPGYRVRLDMLRCIDSGHMSAPLPTCEALSCPELESPP 393
 QY 140 VSGAAEPPOEYFVGVVFEENSGKRIQKQKEMHCSEGLMSNKPQCEVISC-LPPRV 197
 Db 394 HGSMDCSPLRA-FOYDNTCSFRCAGFMALG-ADIVACDNLGOWTAPAPVQALQCODL 451
 QY 198 ENG-D-GIYLKPKYKENERFOYKCKGQFVYKERGDV-CTGSG-WN-POPSCEMTCLTP 252
 Db 452 PVPNARVNCSPFCAFVYQVSCFTCEGILLVG-ACVLOCCLATGNNSVPECOAIPC 510
 QY 253 YIPN-G-IYTPHRI-KHIDEIRKCKNGFRPATRSVSKCTITG-WIPAPR-CSLKPC 307
 Db 511 T-PLSPONGT-MTCVOPLGSSSYKSTCOFTICDEGYSLSGPE-R-IDCTRSGRWTDSP 565
 QY 308 DEPOKRGHRLYEEESRRYFPVPICKKEYSYCDNGFTTPSGSYMDYLCTYNG-WEPEVP 366
 Db 566 MCEAIKCEPLAPPEGSLDCSDTRGEFVNGSTCHSCNNGFRL-BGPNNVCTTSGRGA 624
 QY 367 -C-LRQCIFHYV-EVGESSYVQWR--YLEGSAKYQCHSGSYLPGODTYCTENG-WSP 420
 Db 625 TPPTC 629
 QY 421 -PPKC 424
 RESULT 13
 ID APOH MOUSE STANDARD; PRT: 345 AA.
 AC 001339:
 DT 01-APR-1993 (rel. 25, Last sequence update)
 DT 01-APR-1993 (rel. 25, Last sequence update)
 DT 15-JUL-1999 (rel. 38, Last annotation update)
 DE BETA-2-GLYCOPROTEIN I PRECURSOR (APOLIPOPROTEIN H) (APO-H) (ACTIVATED
 DE PROTEIN C-BINDING PROTEIN) (APC INHIBITOR) (B2GP1).
 GN APOH OR B2GP1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 92372000.
 RA Nonaka M., Matsuda Y., Shiotoishi T., Moriwak K., Natsume-Sakai S.;
 RT "Molecular cloning of mouse beta 2-glycoprotein I and mapping of the
 RT gene to chromosome 11.";
 RL Genomics 13:1082-1087(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX STRAIN-CBA/J; TISSUE-LIVER;
 RX MEDLINE: 94242017.
 RA Seiler G.C., Steel D.M., Zafirooulos A., Seery L.T.,
 RA Whitehead A.S.;
 RT "Characterization, expression and evolution of mouse beta 2-
 RT glycoprotein I (apolipoprotein H).";
 RL Biochem. Biophys. Res. Commun. 200:1521-1528(1994).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX STRAIN-BALB/C; TISSUE-LIVER;
 RA Kristensen T.;
 RT "Structure of the human beta-2-glycoprotein I gene.";
 RT Submitted (FEB-1997) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: BINDS TO VARIOUS KINDS OF NEGATIVELY CHARGED SUBSTANCES
 CC SUCH AS HEPARIN, PHOSPHOLIPIDS, AND DEXTRAN SULFATE. MAY PREVENT
 CC ACTIVATION OF THE INTRINSIC BLOOD COAGULATION CASCADE BY BINDING
 CC TO PHOSPHOLIPIDS ON THE SURFACE OF DAMAGED CELLS.
 CC -1- TISSUE SPECIFICITY: PLASMA.
 CC -1- SIMILARITY: CONTAINS 4 SUSHI (SCR) REPEATS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb.ch/announce/>
 CC or send an email to license@isb.ch).
 CC -----
 CC EMBL: D10056; BAA00945.1;
 DR EMBL: S70439; AAB30789.1;
 DR EMBL: Y11356; CAA72190.1;
 DR PIR: A43286; NBSM-
 DR HSSP: P10998; 1YVC.
 DR MGD: MGI:88058; APOH.
 DR PFM: PFM0084; sushi; 4.
 KW Heparin-binding; Glycoprotein; Plasma; Repeat; Sushi; Signal.
 FT SIGNAL 1 19
 FT CHAIN 20 345
 FT DOMAIN 22 261
 FT REPEAT 22 80
 FT REPEAT 83 138
 FT REPEAT 141 201
 FT REPEAT 204 261
 FT DOMAIN 263 345
 FT DISULFID 23 66
 FT DISULFID 51 79
 FT DISULFID 84 124
 FT DISULFID 110 137
 FT DISULFID 142 188
 FT DISULFID 174 200
 FT DISULFID 205 248
 FT DISULFID 234 260
 FT DISULFID 264 315
 FT DISULFID 300 325
 FT DISULFID 307 345
 FT CARBOHYD 162 162
 FT CARBOHYD 183 183
 FT CARBOHYD 193 193
 FT CONFLICT 252 252
 SQ SEQUENCE 345 AA; 38619 MW; C83F8A6BD51C940 CRC64;
 Query Match 11.1%; Score 375; DB 1; Length 345;
 Best Local Similarity 29.3%; Pred. No. 1,37e-67;
 Matches 73; Conservative 46; Mismatches 115; Indels 15; Gaps 14;
 Db 41 YDPGQIYVSKRPQVSRGKRRFTCLTGMW-PINLRL-CVPRVCFAGILNGIYRYT 98
 QY 24 YSEGOATYKCRPRYKRLGLTVK-VCK-NEWVPSNPSRLCKRRCGHPEDTPPGSRFLA 81
 Db 99 -S-FEYKNTSFACNPGF-PLNGTSSSKTEEGKWSPDIPACARITCPPPYKFKALK 154
 QY 82 VGSFEEFGAKVYTCDEGYQLGEIDRECDADG-WNDIPICEVYKCLPVTLENGRI 140
 Db 155 DYRPSAGNNSLYDVTYFKCLPHFAMIGNDTVMCTEGGNWTRL-PECLEYKCPPPPPEN 213
 QY 141 SGAAEPDOEYFVGVVFEENSGKRIQKQKEMHCSEGLMSNKPQCEVISC-LPPRVEN 199
 Db 214 GYVNPAPKPVLLYDKAFAFGCHETYYKLDGPEAEACTGTGTSFPLTCRE-SCRLPVKAT 272
 QY 200 GGQIT-LKPVYKENERFOYKCKGQFVYKERGDVCTGSG-WNPOPSCEMTCLTPYPNG 257
 Db 273 VLYQGMRYK 281
 QY 256 I-YTPHRIK 265
 RESULT 14
 ID LEM3 RAT STANDARD; PRT: 768 AA.
 AC P98106;
 DT 01-FEB-1996 (rel. 33, Created)
 DT 01-FEB-1996 (rel. 33, Last sequence update)
 DT 15-JUL-1999 (rel. 38, Last annotation update)
 DE P-SELECTIN PRECURSOR (GRANULE MEMBRANE PROTEIN 140) (GMP-140) (PADGEM)
 DE (CD62P) (LEUKOCYTE-ENDOTHELIAL CELL ADHESION MOLECULE 3) (LECAM3).
 GN SELP.

	FT	DISULFID	478	505	BY SIMILARITY
FT	DISULFID	510	554	BY SIMILARITY	
FT	DISULFID	540	567	BY SIMILARITY	
FT	DISULFID	580	624	BY SIMILARITY	
FT	DISULFID	610	637	BY SIMILARITY	
FT	DISULFID	642	686	BY SIMILARITY	
FT	DISULFID	672	699	BY SIMILARITY	
FT	CARBOHYD	45	45	POTENTIAL	
FT	CARBOHYD	54	54	POTENTIAL	
FT	CARBOHYD	107	107	POTENTIAL	
FT	CARBOHYD	212	212	POTENTIAL	
FT	CARBOHYD	347	347	POTENTIAL	
FT	CARBOHYD	456	456	POTENTIAL	
FT	CARBOHYD	603	603	POTENTIAL	
FT	CARBOHYD	654	654	POTENTIAL	
FT	CARBOHYD	661	661	POTENTIAL	
FT	CARBOHYD	679	679	POTENTIAL	
FT	SITE	756	759	POTENTIAL	
SQ	SEQUENCE	768 AA;	83517 MW;	26FD7E8A5F3E1316 CRC64;	
Query Match		11.1%;	Score 374;	DB 1;	Length 768;
Best Local Similarity		26.4%;	Pred. No. 2,51e-67;		
Matches	112;	Conservative	87;	Mismatches 180;	Indels 45; Gaps 43;
Db	290	FECQPEYMRGSDIIHCTDSGOW--SEPLPTCEAIAC-EPLSEPLHSGMCFPSTGAFGY	346		
OY	32	YKCRPEYRFLGTIVVCKN-GEMVSPNSRICRRKPCGHPDTPF-GSFR-LANGSFEF	88		
Db	347	NSSCFERCEGFLVLMGN-DATHCADLIGWTPARVCEALQCQEFPPVSKAO-VS-CSDPF	403		
OY	89	GAKVYITCDEGYQLIGELIDYRECDADG-WINDIPICEVVKCLPTELENGRIVSGAEPD	147		
Db	404	GFLKYQASCSFSCDEGSLVGVASYIRCLATSHMEAPPECOAVSCPTLLSPENGTMCIQ	463		
OY	148	OEYVGGVAVREFCNSNGFKIEGQKEMHCSENGLMWENKRPQCEISCLP-RVENDDGIYK	206		
Db	464	PLGHSNYSSTQCFMDEEF-YLSPERLDCSPSGHWTGSPPMCEALCPETFAPEQGSID	522		
OY	207	PY-YKENERF-QYKCKQCFEYKERGDV-CGSG-WN-POPSEEMCTLPYIPN-GIYT	260		
Db	523	CSHVGEFVSSTGSHFCSEFEELIGSRN-VE-CTVSGRMSAPPPTCKGYSILPVSVRC	580		
OY	261	PH-RIKRIIDELRIEKKNGF-YPAIRSPYKTIIG-W-ITAPRC-SLKPDPOFKH	314		
Db	561	PALTTPGGGTMSCHHLESGEPNTTCYFGKGTFTLRG-A-NSLRGASGOWTAAVPC	637		
OY	315	GRLYX-EESRRPY-FPV-PIKERS-YY-CDNGTTPSGQMYDLRQTVNG-WEPEVP-C	367		
Db	638	RAVKSELHMDTAAVMCSNPMWG-FEYSGTCAFHCPEGSL-NGSARTTCGGDGWSDA	695		
OY	368	LROCI-FHY-VEIG-BSSY-WQRRYIEGSAKYQCHSGISLPNGODIYYCTENG-WSPD	421		
Db	696	MPIC 699			
OY	422	PKC 424			
RESULT	15				
ID	AP0H_HUMAN	STANDARD;	PRT;	345	AA.
AC	P02749;				
DT	21-JUL-1986	(Rel. 01, Created)			
DT	01-JUN-1994	(Rel. 29, last sequence update)			
DT	15-FEB-2000	(Rel. 39, last annotation update)			
DE	BEA-2-GLYCOPROTEIN 1 PRECURSOR (ABOLIPROTEIN H) (APO-H) (ACTIVATED				
DE	PROTEIN C-BINDING PROTEIN) (APC INHIBITOR) (B2GP1).				
GN	AP0H OR B2G1.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;				
OC	Eutheria; Primates; Catarrhini; Homiidae; Homo.				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=LIVER;				
XX	MEDLINE: 91315408.				


```

FT DISULFID 264 315
FT DISULFID 300 325
FT DISULFID 307 345
FT CARBOHYD 149 149
FT CARBOHYD 162 162
FT CARBOHYD 183 183
FT CARBOHYD 193 193
FT CARBOHYD 253 253
FT CARBOHYD 107 107
FT VARIANT 266 266
FT VARIANT 325 325
FT VARIANT 335 335
FT VARIANT 335 335
FT CONFLICT 121 121
FT CONFLICT 188 188
FT CONFLICT 188 188
SO SEQUENCE 345 AA; 38298 MM; 63101704F8EDEF3F CRC64;

```

Query Match 10.9%; Score 367; DB 1; Length 345;
 Best Local Similarity 28.5%; Pred. No. 1.79e-65;
 Matches 72; Conservative 54; Mismatches 106; Indels 21; Gaps 19;

```

DB 40 FYERGEIITSCRGYRGMRKFCPLTGLW-PINTLK-CTPRVCPFGILENGAVRY 97
QY 23 LISEGTOATYKCRPGYRTLTGTYK-VCK-NGEWPSPNSRICKRRKPGHPRDTPFGSFR 80
DB 98 T-T-T-FEYPTNTISFSCNTGTYLNGA-DSAKTEBGMSPELPVCAPICPPPSIPTATL 153
QY 81 AVGSEFEFGAKVYVYTODEGYQLGEIDYRECDADG-WTNDIPICEVYKCLP-VTELENG 137
DB 154 RYKTPSAG-NNSLYRDTAV-FECLPQHAFGNDITICTHGMNTRL-PECREVKCPPSR 210
QY 138 RIVSGAEPDQYEFYFGQVRFECNSGFKIEGKEMHCSENGLWSENKRPQCEVISC-LPR 196
DB 211 PDNGFVYPAKFTLY-YKDKATFGCHDGYSLDGPBELECTKLGNSMSPSCA-SCKVPV 268
QY 197 VENGDDGIY-LKP-VYKENERFQYKCKGQFYKRGDAVCTGSG-WNPQSCCEMTCLTPY 253
DB 269 KKATVYQGERVK 281
QY 254 I-PNGIYTPHRIK 265

```

Search completed: Thu Jun 8 21:50:49 2000
 Job time : 19 secs.

 RELEASE
 (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm
 Run on: Thu Jun 8 21:51:06 2000; Maspar time 36.70 Seconds
 808.665 Million cell updates/sec
 Tabular output not generated.

Title: >US-09-316-163-14
 Description: (1-428) from US09316163.pep
 Perfect Score: 3371
 Sequence: 1 EDCGPPRENSILSGSMS.....DTYCTENGMPKCRVRIK 428

Scoring table:
 Gap 11
 PAM 150

Searched: 225878 seqs, 69334122 residues
 Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database: sptremblj2
 1:sp-archaea 2:sp-bacteria 3:sp-fungi 4:sp-human
 5:sp-invertebrate 6:sp-mammal 7:sp-mhc 8:sp-organelle
 9:sp-phage 10:sp-plant 11:sp-rodent 12:sp-unclassified
 13:sp-vertebrate 14:sp-virus

Statistics: Mean 45.569; Variance 69.358; scale 0.657
 Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	2395	71.0	449	4	014570	COMPLEMENT FACTOR H PR	0.00e+00
2	1720	51.0	669	6	028085	CCP MODULES 3-12, WITH	0.00e+00
3	1136	33.7	452	11	061407	COMPLEMENT FACTOR H-RE	3.25e-267
4	1129	33.5	303	11	061405	COMPLEMENT FACTOR H-RE	2.70e-265
5	790	23.4	808	11	061408	COMPLEMENT FACTOR H-RE	4.80e-173
6	683	20.3	1053	13	091375	COMPLEMENT REGULATORY	2.44e-144
7	520	15.4	533	11	008569	ACROSOMAL MATRIX COMPO	3.02e-101
8	480	14.2	522	6	028769	COMPLEMENT RECEPTOR (F	7.77e-91
9	471	14.0	559	11	062335	COMPLEMENT REGULATORY	1.66e-88
10	470	13.9	657	4	014006	COMPLEMENT H FACTOR (F	3.02e-88
11	470	13.9	679	11	099254	COMPLEMENT RECEPTOR TY	1.25e-84
12	456	13.5	1911	6	029528	COMPLEMENT RECEPTOR 1	1.44e-82
13	448	13.3	331	4	092456	COMPLEMENT FACTOR H-RE	2.60e-82
14	447	13.3	343	11	061406	COMPLEMENT FACTOR H-RE	7.94e-83
15	449	13.3	661	6	029531	COMPLEMENT RECEPTOR 1	4.70e-82
16	449	13.3	2014	6	029530	COMPLEMENT RECEPTOR 1	4.70e-82
17	446	13.2	2039	4	016745	COMPLEMENT RECEPTOR 1	3.53e-77
18	445	13.2	2489	4	016744	UNKNOWN PROTEIN (FRAGM	6.37e-77
19	427	12.7	522	6	028797	FER-2 PRECURSOR.	
20	426	12.6	243	4	014310		

ALIGNMENTS

RESULT ID	1	PRELIMINARY:	PRT:	449 AA.
AC	014570; P78435;			
DT	01-NOV-1996 (TREMBLrel. 01, Created)			
DT	01-NOV-1996 (TREMBLrel. 01, Last sequence update)			
DT	01-NOV-1999 (TREMBLrel. 12, Last annotation update)			
DE	COMPLEMENT FACTOR H PRECURSOR.			
GN	HF OR CFH.			
OS	Homo sapiens (human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;			
CC	Eutheria; Primates; Catarrhini; Homiidae; Homo.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RK	MEDLINE; 88134059.			
RA	RIPACHE J., DAY A.J., HARRIS T.J.R., SIM R.B.;			
RT	"The complete amino acid sequence of human complement factor H.";			
RL	Biochem. J. 249:593-602(1988).			
RN	[2]			
RP	SEQUENCE OF 226-449 FROM N.A.			
RK	MEDLINE; 86169701.			
RA	KRISTENSEN T., WEISEL R.A., TACK B.F.;			
RT	"Structural analysis of human complement protein H: homology with Cdb			
RL	binding protein, beta 2-glycoprotein I, and the Ba fragment of B2.";			
RN	J. Immunol. 136:3407-3411(1986).			
RP	[3]			
RA	SEQUENCE OF 1-19 FROM N.A.			
RL	VIR D.P., WILLIAMS S.A.;			
RT	Submitted (Apr-1996) to the EMBL/GenBank/DBJ databases.			
RN	[4]			
RP	SEQUENCE OF 1-9 FROM N.A.			
RA	DOMINGUEZ O.;			
RL	Thesis (1993), Immunologia, Hospital Trias I Pujol, Spain.			
DR	EMBL; A07523; CA330403.1; -			
DR	EMBL; M12383; AAAS2013.1; -			
DR	EMBL; U56979; AAB01987.1; -			
DR	EMBL; Z29665; CA82763.1; -			
DR	HSSP; P10998; IYVC.			
DR	PFAM; PF00084; susbl; 7.			
KM	Signal.			
FT	SIGNAL.			
FT	CHAIN			
SO	SEQUENCE			
Query Match	449 AA;	51033 MW;	61231E1B CRC32;	
	71.0%;	Score 2395;	DB 4;	Length 449;

21 426 12.6 560 5 022328 COSMID T07H6. 6.37e-77
 22 423 12.5 645 14 09WRU2 COMPLEMENT BINDING PRO 3.73e-76
 23 419 12.4 315 6 028770 COMPLEMENT RECEPTOR (F 3.93e-75
 24 411 12.2 579 11 060736 ZONA PELLUCIDA 3 RECP 4.32e-73
 25 410 12.2 754 6 028290 CELL ADHESION MOLECULE 7.77e-73
 26 409 12.1 497 11 063612 S12 ANTIGEN PRECURSOR. 1.40e-72
 27 404 12.0 550 14 040912 ORF 04. 2.62e-71
 28 400 11.9 550 11 035520 S12 ANTIGEN (FRAGMENT) 2.72e-70
 29 400 11.9 550 14 040912 ORF 4. 2.72e-70
 30 398 11.8 259 14 088903 41KB FRAGMENT FROM LE 8.76e-70
 31 389 11.5 974 5 P91658 FURROEM. 1.68e-67
 32 363 10.8 263 14 089859 HOMOLOG OF VACCINIA VI 5.99e-61
 33 363 10.8 263 14 007033 D12L PROTEIN. 5.99e-61
 34 356 10.6 263 14 089061 B19L. 3.40e-59
 35 356 10.6 263 14 089076 D1780M13.1.2 (SELECTIN 6.03e-59
 36 359 10.6 740 4 095508 MEMBRANE COFACTOR PROT 6.05e-59
 37 355 10.5 377 6 062838 MEMBRANE COFACTOR PROT 6.05e-59
 38 355 10.5 378 6 062837 MEMBRANE COFACTOR PROT 6.05e-59
 39 352 10.4 740 4 095507 D1780M13.1.1 (SELECTIN 6.02e-57
 40 347 10.3 285 6 019121 MEMBRANE COFACTOR PROT 3.36e-56
 41 344 10.2 285 6 019127 MEMBRANE COFACTOR PROT 3.36e-56
 42 344 10.2 285 6 019126 MEMBRANE COFACTOR PROT 3.36e-56
 43 345 10.2 349 4 015429 CD46. 1.90e-56
 44 341 10.1 369 6 P79138 MEMBRANE COFACTOR PROT 1.88e-55
 45 334 9.9 483 11 064735 COMPLEMENT RECEPTOR RE 1.03e-53

Db 3 GSPHAEENQENYKAKVYTTDECGYOMNGEMNFREDCTNGTNDIPICEVVKCLPYTEPE 62
 |||||
 QY 76 GSFRLAVSGSEFPAKRVYITDDEGTQLLGELEIDIRCDADGNTNDIPICEVVKCLPYTELE 135
 |||||
 Db 63 NGKLEPSDLLEDQRYLYQVYVQFECNSSTYMLDGPRLQIHCSAGVWSADTPACVIEIFCKPP 122
 |||||
 QY 136 NGRLVSGAAGPDQRYTGQVYVRFENSSQFKLEGGQKEHMSCSNGLMSNEKPPCVIISCLPP 195
 |||||

OC	RESULT	4	PRELIMINARY;	PRT;	303 AA.
OC	ID	061405			
AC	061405;				
DT	01-NOV-1996 (TREMBLrel.. 01, Created)				
DT	01-NOV-1996 (TREMBLrel.. 01, Last sequence update)				
DT	01-NOV-1999 (TREMBLrel.. 12, Last annotation update)				
DE	COMPLEMENT FACTOR H-RELATED PROTEIN.				
OS	Mus musculus (mouse).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia				

OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 9013969.
 RA VIK D.P., MUNOZ-CANOVES P., KOZONO H., MARTIN L.G., TACK B.F.,
 RA CHAPLIN D.D.;
 RT "Identification and sequence analysis of four complement factor H-
 RT related transcripts in mouse liver."
 RL J. Biol. Chem. 265:3193-3201(1990).
 DR EMBL; M29007; AAA37413.1; -
 DR HSSP; P10998; 1YVC.
 DR PIRAM; PF00084; sush1; 13.
 DR PIRAM; PF00084; sush1; 13.
 SO SEQUENCE 303 AA; 34498 MW; 39350FD1 CRC32;

Query Match 33.5%; Score 1129; DB 11; Length 303;
 Best Local Similarity 70.6%; Pred. No. 2,70e-265;
 Matches 137; Conservative 24; Mismatches 31; Indels 2; Gaps 2;
 Db 15 TAWLSTAKEETKSPYTLNGITPHRIHRSDEIRKCYGFPYTGSTVSKCTPTG 74
 QY 236 SGWNPQSPCEEMTCLTPYIPNGITPHRIHRSDEIRKCYGFPYTGSTVSKCTITG 295
 Db 75 WIPVPCITLPCFPOFKYGRLYEESLRPNFVSIKNGYKCDNGFSPPS-GTSMYTL 133
 QY 296 WIPAPRCSLPCDFPOFKHGRLYEESRRYPVPDYGKSYCDNGFTPSQSY-WDYL 354
 Db 134 RCTAGMEVEVCVRKCYHYENGSDTYMEKIVYOGSLKYQYXSLONGODTMTCT 193
 QY 355 RCTVNMWEVEVCLRCRCHIVHYEGESSYWKRYIEGOSAKYQCHSGISLPGODTYCT 414
 Db 194 ENGMSPPKICIRIN 207
 QY 415 ENGMSPPKICIRIN 428

RESULT 5
 ID 061408 PRELIMINARY; PRT; 808 AA.
 AC 061408;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
 DE COMPLEMENT FACTOR H-RELATED PROTEIN.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 9013969.
 RA VIK D.P., MUNOZ-CANOVES P., KOZONO H., MARTIN L.G., TACK B.F.,
 RA CHAPLIN D.D.;
 RT "Identification and sequence analysis of four complement factor H-
 RT related transcripts in mouse liver."
 RL J. Biol. Chem. 265:3193-3201(1990).
 DR EMBL; M29009; AAA37416.1; -
 DR HSSP; P08603; 1HCC.
 DR PIRAM; PF00084; sush1; 13.
 DR PIRAM; PF00084; sush1; 13.
 SO SEQUENCE 808 AA; 91654 MW; 23644778 CRC32;

Query Match 23.4%; Score 790; DB 11; Length 808;
 Best Local Similarity 66.2%; Pred. No. 4.80e-173;
 Matches 90; Conservative 24; Mismatches 22; Indels 0; Gaps 0;
 Db 14 LTAWLSTAKEGVKSCFEPQFYKGRLYEELIRPNFVSIKNGYKCDNGFSPPSGLEWD 73
 QY 293 ITGWIPAPRCSLPCDFPOFKHGRLYEESRRYPVPDYGKSYCDNGFTPSQSYWD 352
 Db 74 YLRCTVGMKEVEVCVRKCYHYENGSDTYMEKIVYOGSLKYQYXSLONGODTMT 133
 QY 353 YLRCTVNMWEVEVCLRCRCHIVHYEGESSYWKRYIEGOSAKYQCHSGISLPGODTY 412
 Db 134 CTENGMSPPKICIRIK 149
 QY 413 CTENGMSPPKICIRIK 428

RESULT 6
 ID 091275 PRELIMINARY; PRT; 1053 AA.
 AC 091275;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
 DE COMPLEMENT REGULATORY PLASMA PROTEIN.
 OS Parabrax nebulifer (barred sand bass).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
 OC Neopterygii; Teleostei; Euteleostei; Acanthopterygii; Percomorpha;
 OC Perciformes; Percoidae; Serranidae; Parabrax.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX TISSUE-LIVER;
 RX MEDLINE; 94318039.
 RA DAHMEN A., KAIJOH T., ZIPFEL P.F., GIGLI I.;
 RT "Cloning and characterization of a cDNA representing a putative
 RT complement-regulatory plasma protein from barred sand bass (Parabrax
 RT nebulifer)."
 RL Biochem. J. 301:391-397(1994).
 DR EMBL; L21703; AAA92556.1; -
 DR HSSP; P08603; 1HFI.
 DR PIRAM; PF00084; sush1; 16.
 DR PIRAM; PF00084; sush1; 16.
 SO SEQUENCE 1053 AA; 117597 MW; 0DF68EDB CRC32;

Query Match 20.3%; Score 683; DB 13; Length 1053;
 Best Local Similarity 31.7%; Pred. No. 2.44e-144;
 Matches 132; Conservative 76; Mismatches 180; Indels 28; Gaps 22;
 Db 50 YPGGRQYRVGNCNGY-S-GFFKLYCVCAGK-ETRGAK-CQPRSCGHPGQAQFADFLAEG 105
 QY 24 YSEGTQATYKCRPGRYTLGTIVKCKNGEWPNSPSKRRRCGHPGQDTPFSFRLANG 83
 Db 106 NDFVFGKVVYTCQGYOMYRIRYRCVAEGMDGVVPCVESQOC-PLIHVDNNVYIGG 164
 QY 84 SEFERGAKVYTCDEGQYQGLGEIDYRECDDAGWINDIPCEVYKCLPVTLENGRATYSGA 143
 Db 165 --PEBAT-FGNVVRFSKSRSEILDSPELYCDERDMSGVPKCAITCAIPIENGWY 221
 QY 144 AEPDEYVFGGVVFECSNGFKI-EGQKEMHCSENGIMSEKPCQVCSICLIPRVENGDG 202
 Db 222 PGALREYKENVLYEEDRAKHHIDRSTICQKIKAKMSPTPLCESIKRRLTIMDSTRY 281
 QY 203 YLRKPYKENERFYKCGQGVYKREDA-VCTG--SGWNPQSPCEEMTCLTPYIPNGIT 259
 Db 282 EPAYNLPSPGELTKVICARTSWISTPOETSVTTCODNGEWSIRPTCOEYKCSNRREP 341
 QY 260 TP-HAIKIRIDDELRYEC-KNGFYPATR-SPV-SKCTITG-WIPAPRCSLPCDFPOFKH 314
 Db 342 VDSWDVRSWERY--TLDNTRWYCKRGYKRTGYTMA--TCGNGMAMPNLECYKCSK 396
 QY 315 GRLYEESRRYPVPDYGKSYCDNGFTPSQSYWDYLRCTVNGEPEVPC-LRCOIF 373
 Db 397 ENIGDAVIVGDQKQYLNOKAIVACGEG-N--RGRITLTGCGMSGDRCT-VK 448
 QY 374 HYVEYEGESSYWR-R-YIEGOSAKYQCHSGISLPGODTYCTENGMSPPKICIRIK 428

RESULT 7
 ID 008569 PRELIMINARY; PRT; 533 AA.
 AC 008569;
 DT 01-JUL-1997 (TREMBLrel. 04, Created)
 DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
 DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
 DE ACROSOMAL MATRIX COMPONENT AM67 PRECURSOR.
 OS Cavia porcellus (Guinea pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Rodentia; Hystriocognathi; Cavidae; Cavia.
 RN [1]
 RP SEQUENCE FROM N.A.
 RP STRAIN-HARTLEY; TISSUE-TESTIS;

RX MEDLINE: 97284752.
 RA FOSTER J.A., FRIDAY B.B., MAULIT M.T., BLOBEL C., WINFREY V.P.,
 RA OLSON G.E., KIM K.S., GERTON G.L.;
 RT "Am67, a secretory component of the guinea pig sperm acrosomal matrix,
 RT is related to mouse sperm protein sp56 and the complement component 4,
 RT binding proteins.";
 RL J. Biol. Chem. 272:12714-12722(1997).
 DR EMBL: U75654; AAC13888.1;
 DR HSSP: P10998; IYVC.
 DR PFAM: PF00084; sushi; 7.
 KW Signal.
 FT CHAIN 1 28 POTENTIAL.
 FT SEQUENCE 533 AA; 59772 MW; 60E5A526 CRC32;
 SQ
 Query Match 15.4%; Score 520; DB 11; Length 533;
 Best Local Similarity 27.2%; Pred. No. 3.02e-101;
 Matches 123; Conservative 105; Mismatches 179; Indels 46; Gaps 37;
 Db 29 DC-GPPILPFPASVLYQSY-ETNFTGTALKYNGHGYRWYSSHVLDINGSWI-VN-V 84
 QY 2 DCKPPPPREN-SEILSGMSSEQLYSEGTQATYKCRPGYRTGLTIVKVC-NGEWPSNPS 59
 Db 85 -FCAKKCGRNGLANG--KWEITDLFGSTIEFSCSKGSLIGSTT-SCCESQGTVD 140
 QY 60 RICKRRGCHGHDTPFGSFRILAVGSEFEFGAKVYTCDEGYQLLEIDYRCDADG--- 115
 Db 141 WSDLPFCVYIKCDSPDISNGK-HSGTDE-DL-YTGSGLTYVCDPYSLLGNASISCL 197
 QY 116 WTNDPICEVYKCLPVELENGRIVSGAEPDQYFYFGQVYFECNSGFKLEGKEMHC- 174
 Db 198 VAKTVGVWSSNPPTCEKVICRPHIKRGLFLSGFGFYTKTLVTSCKKGYLLRGSSI 257
 QY 175 -SEN-GLMWENKQCVIEISCLPRVENGDIY-LKPYKENERFOYKCKGQFYKRGD 230
 Db 258 IHCEANSKWPSIPTCPNCGCIDPEVYIYSWENVLISLNOELFEGLSLKYCKGIR 317
 QY 231 AVC-TSGGNPQ-PSCEMTC--LT--PIYD-N-GIYT-PHRIKHRIIDDETRCKNGFY 281
 Db 318 PTNEPPTVTCQENLKAISKGERVCCPTPMMEKRII-NE-RRDPTGVVAYEYIF 375
 QY 282 PATRSPVS-KCITT-GWIPAPRCSLKPCDPQFKHSLYEESSRPFPYPI-G-KEYSY 337
 Db 376 YMCDEGYPIISADGRS--SQADGMMNPKPACESANCLPDLINGLSLYEKDIYETEN 433
 QY 338 Y-CDNGFTTSSQSYWDYLRCTVNG-WPEVP-CLRQ-CIFHYEYGSYSWMORRYEGQS 393
 Db 434 VTIHCDGYEVGPONII-CSENRTWPEIIPKC 465
 QY 394 AKVOCHSGSYSLPNGQDYYCTEN-GMSPP-PKC 424
 RESULT 8
 ID 028769 PRELIMINARY; PRT; 522 AA.
 AC 028769;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DT 01-NOV-1996 (TREMblrel. 12, Last annotation update)
 DE COMPLEMENT RECEPTOR (FRAGMENT).
 OS Papilio hamadryas (Hamadryas baboon).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Cercopithecoidea; Cercopithecoinae;
 OC Papio.
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE-BONE MARROW;
 RA BIRNICHAM D.J., LOGAR C.M., SHEN X.P., CHEN W.;
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U79797; AAA99004.1;
 DR HSSP: P08603; IHFI.
 DR PFAM: PF00084; sushi; 7.
 DR NON_TER 1
 FT SEQUENCE 522 AA; 56626 MW; 7862072C CRC32;
 SQ

Query Match 14.2%; Score 480; DB 6; Length 522;
 Best Local Similarity 33.2%; Pred. No. 7.77e-91;
 Matches 107; Conservative 74; Mismatches 102; Indels 39; Gaps 33;
 Db 68 FPGVLYKCYELLPCHG-KPSITICLKNSVTSK-DK-CTRKSCRNPKD-PVNGM-VHV 122
 QY 24 YSEGTQATYKCRPGYRTGLTIVKVC-KNGEWPSNPSICKRRKRGHDPFGSFRILAV 82
 Db 123 IKDIQFSQINYSCKNGRILGSSSACIIISGNTVIMDNERTICEIIPGCPPT-INGND 181
 QY 83 GSEFEFGAKVYITODEGYQLLEIDYR-ECDADG-WTNDPICEVYKCLPVELENGR 138
 Db 182 FIS-TS--RETPYGSVYTRCNLGSGRKKLFELVSEPSITYSKDDQYIGSGAPQC 237
 QY 139 IVSGAEPDQY-FGQVYFECN--SGFK-I-E--GQKEMHC-S-EN-GLMWENKQCV 187
 Db 238 IIPKCMPPVNEGVLSVNSLSLMEVVEFRCPQPFYKGPVHVOQALNKEPELPS 297
 QY 188 VEIS-CLPRVENGDIY-LKPYKENERFOYKCKGQFYKRGDAVCTG-SGMNP-PS 243
 Db 298 CSRY-COPPEILGHTPSSHODFSQGEVYFSCPEG-Y-DLRGAASLHCTPOGWNPEA 354
 QY 244 CEEMTCLTP-YIPRGVYTPHRIKHRIIDDEIRYCKNGFYPATSPVS-KCITIG-WIP-A 239
 Db 355 PICVYKSCDDFLQOLPHGRVLF 376
 QY 300 PRCSLKPCD-F-PQFKHGRLYY 319
 RESULT 9
 ID 063135 PRELIMINARY; PRT; 559 AA.
 AC 063135;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DT 01-NOV-1996 (TREMblrel. 12, Last annotation update)
 DE COMPLEMENT REGULATORY PROTEIN.
 GN CRRY.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN-SPRAGUE-DAWLEY;
 RA MEDLINE: 96006570.
 RA QUIG R.J., LO C.F., ALEXANDER J.J., SNEED A.E., MOXLEY G. III;
 RT "Molecular characterization of rat Crry: widespread distribution of
 RT two alternative forms of Crry mRNA."
 RL Immunogenetics 42:362-367(1995).
 RN (2)
 RP SEQUENCE FROM N.A.
 RC TISSUE-LIVER;
 RA DOHI N., SAKURADA C., NONAKA M., OKADA N., OKADA H.;
 RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U36532; AAA91821.1;
 DR EMBL: D42115; BAA22548.1;
 DR HSSP: P10998; IYVC.
 DR PFAM: PF00084; sushi; 7.
 FT SEQUENCE 559 AA; 61680 MW; C87EEC58 CRC32;
 SQ
 Query Match 14.0%; Score 471; DB 11; Length 559;
 Best Local Similarity 26.9%; Pred. No. 1.66e-88;
 Matches 123; Conservative 114; Mismatches 170; Indels 50; Gaps 36;
 Db 37 OCPAPPLPFPYAKINPT-DESNFPGTSLKYRCRGYIKRQPSIT-CEVNSVWT-S-POD 92
 QY 2 DCKGPPPPRENSEILSGMSSEQLYSEGTQATYKCRPGYRTGLTIVKVC-NGEWPSNPSR 60
 Db 93 VCIRKQCEPLDPONG-I-VHVNIDIRFGSSITTYCNGNGRYLIGSSAMCIISDQSVAMD 150
 QY 61 ICRKRPCGHDTPFGSFRILAVGSEFEFGAKVYTCDEGYQLLEIDY--YECADG-GWT 117
 Db 151 AEAFCESIPCEIIPSIINGDFES--PNRE-DFHYGVVYVQCNTDARGKFLNLVGEPS 207

```

QY 118 NDICEVYKCLPYTELENGRIVSGAAPPDDEYFGQVRECN--G---FKIEGKE 171
DB 208 ICHTSIDQVWVGSPPEQCIENKCTPPHVENAVIVKNSLFLRDMVEFRODGFMM 267
QY 172 MHC-S-EN-GLMSNKKQCVIEIS-CLPFRYENGITL-KRYKENERFOYKCKOGVY 225
DB 268 KGDSSVYCRSINRREPOLPSCFKVKSAGALGELPNHVFVQNL-Q-LGAKVTFVQNTG 325
QY 226 KERDAVCTG-SGWNPO-PSC-EEWTC--LTPYIPNG-IYMPHRIKHRIIDEIRECKNG 279
DB 326 -YQKGNSSSHCVLDGVSINSSVPCQVYICKLPQMSGFQKQKQDY--YQDN 381
QY 280 FYPTSPVSKCTITG---WIPA-PROSLKPCDFPQFKHRLYYEESRRPFPVPIKE 334
DB 382 VALCEDEGTYLEGGSSQSC-OSDAS-WDPLPKVSYQVICKLPQMSGFQKQKQDY 439
QY 335 YSYCDNNGFTTPOSYNDYLACTYNGMEPEVP-CLRQ--CLFHYEYG-ESSYQORR-YI 389
DB 440 YQDNVALECEDEGTYLEGGSSQSCSDASWDPPLPKCV 476
QY 390 EGOSAKVQCHSGYSLPNQDITYCTENGWSP-PCV 425

RESULT 10
ID 014006 PRELIMINARY: PRT: 657 AA.
AC 014006;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-NOV-1999 (TREMblrel. 12, Last annotation update)
DE COMPLEMENT H FACTOR (FRAGMENT).
GN HP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Homiidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 88025472.
RA DAY A.J., RIPOCHE J., LYONS A., MCINTOSH B., HARRIS T.J., SIM R.B.;
RT "Sequence analysis of a cDNA clone encoding the C-terminal end of
RT human complement factor H."
RL Biosci. Rep. 7:201-207(1987).
DR EMBL: M17517; AAA52016.1; -.
DR HSSP: P08603; 1HFH.
DR PFAM: PF00084; sushi; 11.
DR NON_TER 1
SQ SEQUENCE 657 AA; 74247 MW; F4AB5238 CRC32;

Query Match 13.9%; Score 470; DB 4; Length 657;
Best Local Similarity 26.0%; Pred. No. 3.02e-88;
Matches 107; Conservative 78; Mismatches 201; Indels 26; Gaps 25;

DB 197 FDHNSNIRYROR-G-KE-GWHTVINGRMDPEVNCMAQIOLCPPPOLP-NSHNMTTT 252
QY 24 YSECTQATYRCRPEYRLGLTIVKCKNGEWPNPSRICKRRPGHGDTPFGSFRALV 83
DB 253 LNYDEKRYVLCQENT-LIOEGEITCKDGRW-OSIPLCVEKIKPSQPPQIEHGTINS 310
QY 84 SEFEFGAKVYVTCDEGYQLGEIDYRECDADGWTNDIPIC-EYVCKLPVTELENGRIYSG 142
DB 311 RS-SQESYAGTKLSTYCEGGFRISENETTCY-MGKMSSP-POCEGLPKCSPEISGV 367
QY 143 AAEPDQYTYGQVYRFBECNSGFKIEGKEMHCSENGLSMNSKPKCVESLCL-PRVENG 201
DB 368 VAHMSDSYQGEVYTKCFEGFIDGPAIAKCLGEMSHSPSCIKTCLSLPSENAIPM 427
QY 202 GYLKPYKENERQYKCKGQFYKRGDAVCTSGMNPDPSCHEMCLT-PYIPNGIYT 260
DB 428 GEKRDYKAAEQVYTC-ATYYKMGDSANTCINSMTGRTCDTSCVNDPIVQNA-YI 485
QY 261 PHRTK-HRIDEREKCKNGFPATRSVSKCTITGWIIPARCSLKPDPFOFKHRLY 319
DB 486 VSRQMSKYP-S-GERVYQCRSPYEMFGDE--EYV-CLNGMWTPEPQCKSTGKCPFP 540

```

```

QY 320 EESRRPYPPVPIGKREYCYCDNGFTTPOSYNDYLCTVNGW-EP-EV-PCLRQC-IFHY 375
DB 541 IDNGITTFPLSYAPASSVYQCONLYQL-EGNKRTICRNGQWSEPPKCH 591
QY 376 VEYGE-SSYQORRYIEGSAKVQCHSGYSLPNQDITYCTENGWSPPKCV 426

RESULT 11
ID 099254 PRELIMINARY: PRT: 679 AA.
AC 099254;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-NOV-1999 (TREMblrel. 12, Last annotation update)
DE COMPLEMENT RECEPTOR TYPE 2 PRECURSOR (CR2)
DE (ALTERNATIVELY SPLICED VERSION) (FRAGMENT).
GN CR2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-BALB/C.
RX MEDLINE: 90229754.
RA KURTZ C.B., O'TOOLE E., CHRISTENSEN S.M., WEIS J.H.;
RT "The murine complement receptor gene family. IV. Alternative splicing
RT of Cr2 gene transcripts predicts two distinct gene products that share
RT homologous domains with both human CR2 and CRI1."
RL J. Immunol. 144:3581-3591(1990).
RN [2]
RP SEQUENCE OF 21-367 FROM N.A.
RX MEDLINE: 95105691.
RA KIM Y.U., KINOSHITA T., MOLINA H., HOURCADE D., SEYA T., WAGNER L.M.,
RA HOLERS V.M.;
RT "Mouse complement regulatory protein Crry/p65 uses the specific
RT mechanisms of both human decay-accelerating factor and membrane
RT cofactor protein."
RL J. Exp. Med. 181:151-159(1995).
CC -1- ALTERNATIVE PRODUCTS: ALTERNATIVE SPLICING OF CR2 GENE TRANSCRIPTS
CC PREDICTS TWO DISTINCT GENE PRODUCTS THAT SHARE HOMOLOGOUS DOMAINS
CC WITH BOTH HUMAN CR2 AND CRI1.
CC -1- TISSUE SPECIFICITY: B LYMPHOCYTES.
DR EMBL: U17128; AAA78271.1; -.
DR EMBL: U17123; AAA78271.1; JOINED.
DR EMBL: U17124; AAA78271.1; JOINED.
DR EMBL: U17125; AAA78271.1; JOINED.
DR EMBL: U17126; AAA78271.1; JOINED.
DR EMBL: U17127; AAA78271.1; JOINED.
DR EMBL: M36470; AAA37449.1; -.
DR HSSP: P10998; 1VVC.
DR MGD: MGI:88489; Cr2.
DR PFAM: PF00084; sushi; 10.
DR KW Complement pathway; Glycoprotein; Transmembrane; Repeat; Signal;
KW Alternative splicing; Receptor.
FT SIGNAL 1
FT CHAIN 12 679 COMPLEMENT RECEPTOR TYPE 2.
FT NON_TER 679
SQ SEQUENCE 679 AA; 74916 MW; BF08AC75 CRC32;

Query Match 13.9%; Score 470; DB 11; Length 679;
Best Local Similarity 31.7%; Pred. No. 3.02e-88;
Matches 101; Conservative 72; Mismatches 111; Indels 35; Gaps 25;

DB 42 FAITWEYCRPGYFRKSFIT-CLFTSKW--SDAOPCKRRCMPPOEHLGSHVINT 98
QY 24 YSECTQATYRCRPEYRLGLTIVKCK-NGEKWPNPSRICKRRPGHGDTPFGSFRALV 82
DB 99 G-LEFGSTTYSQNOGRRLIGSSATCIYSDNTVMDNDMPLOCESIPCESPAISNGF 156
QY 83 GSEFEFGAKVYVTCDEGYQLGE--IDYRECDADG-WTNDIPICEVYKCLPYTELENGRI 139
DB 157 YS-SS-RD-SFYGVMVYTYCHGKNREKLFDLVGEESIYCTSKDNQVYGLWNSDPPOCIP 213

```

```

QY 140 VSGAABDQEFYFGQVVRFECSNG-----FKISGOKEMHC-S-EN--GLMSNEKPCQVE 189
Db 214 RVCMPMEIENGVLSEGP-KHSFLNDTVIFKCKSGFTMGSRILAKOMPNSKSPLEPTC 272
QY 190 -ISCLPRVENG--DGILYLRKYKENERFOYCKCKGQFYKBERGAVC-TGSGMNPQ-PSC 244
Db 273 F-MCCLPPOINTLHGDKKDEFFSVGOKVSVTCNPG-YTLIGTNLVECTSLGTWNTVPT 330
QY 245 EEMTCLLP-YIPNGIYPRHRIKHIDDEIRIECKANGFYPAIRSVSKCTITG-WIPA-PR 301
Db 331 CEVKSCDAPNHLHGRVF 349
QY 302 CSKPCD-FP-QFKHGRLY 318

RESULT 12
ID Q29528 PRELIMINARY; PRT; 1911 AA.
AC Q29528;
DT 01-NOV-1996 (TREMblrel, 01, Created)
DT 01-NOV-1996 (TREMblrel, 01, Last sequence update)
DT 01-NOV-1999 (TREMblrel, 12, Last annotation update)
DE COMPLEMENT RECEPTOR 1 (FRAGMENT).
GN CRI1.
OS Papio hamadryas (Hamadryas baboon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Cercopithecoidea; Cercopithecinae;
OC Papio.
RN [1].
RP SEQUENCE FROM N.A.
RA CLEMENZA L., SUBRAMANIAN B.V., NICKELLS M.W., HOURCADE D.E.,
RA ATKINSON J.P.;
RL Submit (MAR-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; L39791; AAA62170.1; -.
DR HSSP; P08603; IHCC.
DR PIRAM; PF00084; sushi; 29.
FT NONTER 1911 1911
SQ SEQUENCE 1911 AA; 210173 MW; 03D640C2 CRC32;

Query Match 13.5%; Score 456; DB 6; Length 1911;
Best Local Similarity 31.8%; Pred. No. 1,25e-84;
Matches 118; Conservative 81; Mismatches 128; Indels 44; Gaps 37;

Db 850 FPIGTSLKCYGKREYGC-KPFSITCLD-NLYWSSPKDYCKKSKCTPPD-FVNGM-VHYI 905
QY 24 YSEGTQATYKRCPRGLTGLTYKCKNGEWPNSNRKRCRPGDTPFGSFLAVG 83
Db 906 TDIVGSRINSCCTGHLIGHS--AECIIIGNTAWSTKRPICORICGLP-PIANGD 963
QY 84 SEFERGAVYTCDEGYLLGEIDYREC--DAD-G-WINDIPICVVC-LPYTELENGR 138
Db 964 FIS--T-NREYFHYGVYTYRCNLGSRKRLFELVGEPSIYCTSKDQVIGSGPADQC 1019
QY 139 IVSGAABDQGY-FGQVVRFECSN--SGFK-I-E--GQKEMHC-S-EN--GLMSNEKPCQ 187
Db 1020 IIPKCMPPNENGLVSVNSLSFLNEVEFRQCPGVYMGPRVQCALNKWPELPS 1079
QY 188 VEIS-CLPVRVNGDGYL-KPYKENERFOYCKCKGQFYKBERGAVC-TGSGMNPQ-PS 243
Db 1080 CSRV-COPPELHGHETPSHODKSPSCQEVFYSCPEP-Y-DLRGAALHCTPOQDMPE 1136
QY 244 CEEMTCLLP-YIPNGIYPR-HRIKHRIIDDEIRIECKANGFYPAIRSVSKCTITG-WIPA- 298
Db 1137 APCAVKSCD--DF-LGQILHGRVLP-FNLQLGAKVSFVODEGFRILGSSVSHCVLWGM 1192
QY 299 APRCSLKPCDQFQFKHGRLYEESRRPYFPVPIGKEYSYCDNGFTTSPQSQYMDLACTV 358
Db 1193 RSLMNSNVPVC 1203
QY 359 NG-WEPEVP-C 367

RESULT 13
ID Q29496 PRELIMINARY; PRT; 331 AA.

```

```

AC Q29496;
DT 01-FEB-1997 (TREMblrel, 02, Created)
DT 01-FEB-1997 (TREMblrel, 02, Last sequence update)
DT 01-NOV-1999 (TREMblrel, 12, Last annotation update)
DE COMPLEMENT FACTOR H-RELATED PROTEIN 4 PRECURSOR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominoidea; Homo.
RN [1].
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE; 97190290.
RA SKERA C., HELLMAGE J., WEBER W., TILKORN A., BUCK F., MARTI T.,
RA KAMPE E., BEISIEGEL U., ZIEFEL P.F.;
RT "The human factor H-related protein 4 (FHR-4). A novel short consensus
RT repeat-containing protein is associated with human triglyceride-rich
RT lipoproteins."
RL J. Biol. Chem. 272:5627-5634(1997).
DR EMBL; X98337; CAA66980.1; -.
DR HSSP; P10998; IIVC.
DR PIRAM; PF00084; sushi; 5.
KW Signal.
FT SIGNAL 1 19
FT CHAIN 20 331
SQ SEQUENCE 331 AA; 37325 MW; 2505D66C CRC32;

Query Match 13.3%; Score 448; DB 4; Length 331;
Best Local Similarity 43.6%; Pred. No. 1,44e-82;
Matches 61; Conservative 22; Mismatches 52; Indels 5; Gaps 4;

Db 9 LTLWVSCANGQEVKPCDPEIIOHGLYKSLRLYPFAAGSYSYCDONFVTPSGSYW 68
QY 293 ITGMIPAPRC-SLKPCDQFQFKHGRLYEESRRPYFPVPIGKEYSYCDNGFTTSPQSQY 351
Db 69 DYHCTODGWLTPYPCARTSKSDIEENGFISSSIYLNKRYQKCKPGVATAAGNS 128
QY 352 DYHCTVNGMEPEVPCLRQCT-FHY-VEIGESSYWRRIEIGSAKYCHSGISLPGQD 409
Db 129 SGTITCLONGMSAOPICIKF 148
QY 410 T-YCTENGWSPPKCVRI 427

RESULT 14
ID Q61406 PRELIMINARY; PRT; 343 AA.
AC Q61406;
DT 01-NOV-1996 (TREMblrel, 01, Created)
DT 01-NOV-1996 (TREMblrel, 01, Last sequence update)
DT 01-NOV-1999 (TREMblrel, 12, Last annotation update)
DE COMPLEMENT FACTOR H-RELATED PROTEIN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1].
RP SEQUENCE FROM N.A.
RX MEDLINE; 90153969.
RA VIK D.P., MONOZ-CANOVES P., KOZONO H., MARTIN L.G., TACK B.F.,
RA CHAPLIN D.D.;
RT "Identification and sequence analysis of four complement factor H-
RT related transcripts in mouse liver."
RL J. Biol. Chem. 265:3193-3201(1990).
DR EMBL; M29008; AAA37414.1; -.
DR HSSP; P08603; IHFI.
DR PIRAM; PF00084; sushi; 4.
SQ SEQUENCE 343 AA; 38443 MW; 6CFD5090 CRC32;

Query Match 13.3%; Score 447; DB 11; Length 343;
Best Local Similarity 41.4%; Pred. No. 2,60e-82;
Matches 55; Conservative 29; Mismatches 49; Indels 0; Gaps 0;

Db 14 LTSWFSTAKGEVSLCDPFRKIRHGILYDEKKNPEPSSVGLKILLYSCYNFAPSNSEFWT 73
QY 293 ITGMIPAPRC-SLKPCDQFQFKHGRLYEESRRPYFPVPIGKEYSYCDNGFTTSPQSQYMD 352

```